A Proof of Theorem 2

(i) If the likelihood ratio $p(X|\theta_i)/p(X|\theta_j)$ is an absolutely continuous random variable for any $i \neq j$, then $p(X|\theta_j)$, $j = 1, \ldots, m$, have the same support and the posterior distribution is well defined for any prior π and ν-almost all $x \in \mathcal{X}$. Moreover, ties in the posterior probabilities ($p(\theta_i|X) = p(\theta_j|X)$, $i \neq j$) happen with probability zero under any $\theta \in \Theta$. An HPD credible set $\varphi(\cdot, \cdot; \pi)$ is uniquely defined and continuous in π whenever there are no ties in the posterior probabilities. The function $z(\pi)$ defined in Theorem 1 is therefore continuous in π and Theorem 1 implies that there exists a prior $\pi^\star$ for which $\varphi(\cdot, \cdot; \pi^\star)$ has coverage of at least $1-\alpha$.

(ii) Next, let us show that $\pi^\star_j > 0$ for any j and $\varphi(\cdot, \cdot; \pi^\star)$ is a similar $1-\alpha$ confidence set. If $\pi^\star_j = 0$ for some j then θ_j is not contained in the $1-\alpha$ HPD credible set for any x and $\varphi(\cdot, \cdot; \pi^\star)$ has zero coverage at θ_j. Thus, $\pi^\star_j > 0$ for all j. Since $\varphi(\cdot, \cdot; \pi^\star)$ is a $1-\alpha$ credible set

$$\sum_{j=1}^{m} \varphi(\theta_j, x; \pi^\star) p(x|\theta_j) \pi^\star_j = \alpha \sum_{j=1}^{m} p(x|\theta_j) \pi^\star_j.$$

Integration of the last display implies

$$\sum_{j=1}^{m} \left[\int \varphi(\theta_j, x; \pi^\star) p(x|\theta_j) d\nu(x) \right] \pi^\star_j = \alpha. \tag{12}$$

Since the coverage of $\varphi(\cdot, \cdot; \pi^\star)$ is at least $1-\alpha$, $\int \varphi(\theta_j, x; \pi^\star) p(x|\theta_j) d\nu(x) \leq \alpha$. Because $\pi^\star_j > 0$ for all j the equality in (12) can hold only if $\int \varphi(\theta_j, x; \pi^\star) p(x|\theta_j) d\nu(x) = \alpha$ for all j or, in other words, $\varphi(\cdot, \cdot; \pi^\star)$ is similar.

(iii) It suffices to show that if $\sum_{j=1}^{m} \varphi'(\theta_j, x) \geq \sum_{j=1}^{m} \varphi(\theta_j, x; \pi^\star)$ for all $x \in \mathcal{X}$ and $\sum_{j=1}^{m} \varphi'(\theta_j, x) > \sum_{j=1}^{m} \varphi(\theta_j, x; \pi^\star)$ for all $x \in \mathcal{X}_l$ with $\nu(\mathcal{X}_l) > 0$, then

$\int \varphi'(\theta_j, x)p(x|\theta_j)d\nu(x) > \alpha$ for some j.

The HPD set $\varphi(\theta, x; \pi^\star)$ can be defined for ν-almost all x by the minimum length property that for all φ'' with $\sum_{j=1}^m \varphi''(\theta_j, x) = \sum_{j=1}^m \varphi(\theta_j, x; \pi^\star)$, $\sum_{j=1}^m \varphi(\theta_j, x; \pi^\star)p(\theta_j|x) \leq \sum_{j=1}^m \varphi''(\theta_j, x)p(\theta_j|x)$. Thus, for ν-almost all $x \in \mathcal{X}_l$,

$$\sum_{j=1}^m \varphi(\theta_j, x; \pi^\star)p(x|\theta_j)\pi_j^\star < \sum_{j=1}^m \varphi'(\theta_j, x)p(x|\theta_j)\pi_j^\star$$

and for all $x \in \mathcal{X}$ the inequality holds weakly. Integrating this inequality with respect to ν yields $\sum_{j=1}^m \pi_j^\star \int(\varphi(\theta_j, x; \pi^\star) - \varphi'(\theta_j, x))p(x|\theta_j)d\nu(x) < 0$. Since $\sum_{j=1}^m \pi_j^\star \int \varphi(\theta_j, x; \pi^\star)p(x|\theta_j)d\nu(x) = \alpha$ by part (ii), this implies that there exists j such that $\int \varphi'(\theta_j, x)p(x|\theta_j)d\nu(x) > \alpha$.

B Proof of Theorem 3

(i) By the assumed uniform continuity and $\max_j \operatorname{diam}(\tilde{\Theta}_j^m) \to 0$, there exists M_ϵ such that for any $m \geq M_\epsilon$ and $\tilde{\theta}_1, \tilde{\theta}_2 \in \tilde{\Theta}_j^m$, $j = 1, \ldots, m$,

$$\int |\tilde{p}(x|\tilde{\theta}_1) - \tilde{p}(x|\tilde{\theta}_2)|d\nu(x) < \epsilon. \tag{13}$$

In order to obtain a contradiction, assume there exists j^* and $\tilde{\theta}^* \in \tilde{\Theta}_{j^*}^m$ such that

$$\int \tilde{\varphi}^m(\tilde{\theta}^*, x; \pi^{m\star})\tilde{p}(x|\tilde{\theta}^*)d\nu(x) < \alpha - \epsilon. \tag{14}$$

For any $\tilde{\theta}_1 \in \tilde{\Theta}_{j^*}^m$, $\tilde{\varphi}^m(\tilde{\theta}^*, x; \pi^{m\star}) = \tilde{\varphi}^m(\tilde{\theta}_1, x; \pi^{m\star})$ as $\tilde{\varphi}^m$ is constant on $\tilde{\Theta}_{j^*}^m$ by definition. Therefore, by (13) and (14), $\int \tilde{\varphi}^m(\tilde{\theta}_1, x; \pi^{m\star})\tilde{p}(x|\tilde{\theta}_1)d\nu(x) < \alpha$, $\forall \tilde{\theta}_1 \in \tilde{\Theta}_{j^*}^m$, which would make the equality in (6) impossible. A contradiction for $\int \tilde{\varphi}^m(\tilde{\theta}^*, x; \pi^{m\star})\tilde{p}(x|\tilde{\theta}^*)d\nu(x) > \alpha + \epsilon$ can be obtained in the same way.

(ii) Suppose the claim does not hold. Then, there exists a subsequence $\{m_k\}$ with

$$\int \tilde{\phi}(\tilde{\theta}, x) d\tilde{\theta} \geq (1+\epsilon) \cdot \int \tilde{\varphi}^{m_k}(\tilde{\theta}, x; \pi^{m_k\star}) d\tilde{\theta}$$

for ν-almost all x. Pick $m_k > M_{\alpha\epsilon}$, with M_ϵ defined in part (i) of this proof. For this m_k, and any $\tilde{\theta} \in \tilde{\Theta}_j^{m_k}$, define

$$\phi'(\theta_j, x) = \tilde{\phi}'(\tilde{\theta}, x) = \int_{\tilde{\Theta}_j} \tilde{\phi}(\tilde{\theta}_1, x) d\tilde{\theta}_1 / V_j \text{ and } \phi''(\theta_j, x) = \tilde{\phi}''(\tilde{\theta}, x) = \tilde{\phi}'(\tilde{\theta}, x)/(1+\epsilon),$$

where $V_j = \mathrm{vol}(\tilde{\Theta}_j^{m_k})$, $j = 1, \ldots, m_k$. Note that

$$\int \tilde{\phi}(\tilde{\theta}, x) d\tilde{\theta} = \int \tilde{\phi}'(\tilde{\theta}, x) d\tilde{\theta} = (1+\epsilon) \int \tilde{\phi}''(\tilde{\theta}, x) d\tilde{\theta} = (1+\epsilon) \sum_j \phi''(\theta_j, x) V_j.$$

Thus, $\sum_j \phi''(\theta_j, x) V_j \geq \int \tilde{\varphi}^{m_k}(\tilde{\theta}, x; \pi^{m_k\star}) d\tilde{\theta} = \sum_j \varphi^{m_k}(\theta_j, x; \pi^{m_k\star}) V_j$, and since $\varphi^{m_k}(\theta, x; \pi^{m_k\star})$ maximizes $\sum_j \varphi(\theta_j, x) V_j$ subject to $\sum_j [\alpha - \varphi(\theta_j, x)] p(x|\theta_j) \pi_j^{m_k\star} \leq 0$,

$$\sum_j \phi''(\theta_j, x) p(x|\theta_j) \pi_j^{m_k\star} \geq \sum_j \varphi^{m_k}(\theta_j, x; \pi^{m_k\star}) p(x|\theta_j) \pi_j^{m_k\star}.$$

Integration of this inequality with respect to ν gives

$$\sum_j \pi_j^{m_k\star} \int \phi''(\theta_j, x) p(x|\theta_j) d\nu(x) \geq \alpha.$$

Thus, there exists j^* such that $\int \phi''(\theta_{j^*}, x) p(x|\theta_{j^*}) d\nu(x) \geq \alpha$ and

$$\int \phi'(\theta_{j^*}, x) p(x|\theta_{j^*}) d\nu(x) \geq (1+\epsilon)\alpha. \tag{15}$$

At the same time,

$$\int \phi'(\theta_{j^*}, x) p(x|\theta_{j^*}) d\nu(x) = \int_{\tilde{\Theta}_{j^*}^{m_k}} \int_{\tilde{\Theta}_{j^*}^{m_k}} \int \tilde{\phi}(\tilde{\theta}_1, x) p(x|\tilde{\theta}_2) d\nu(x) d\tilde{\theta}_1 d\tilde{\theta}_2 / V_{j^*}^2.$$

Because $\int \tilde{\phi}(\tilde{\theta}_1, x) p(x|\tilde{\theta}_1) d\nu(x) \leq \alpha$ for all $\tilde{\theta}_1 \in \tilde{\Theta}$ by assumption on $\tilde{\phi}$,

$$\int \phi'(\theta_{j^*}, x) p(x|\theta_{j^*}) d\nu(x) \leq \alpha + \sup_{\tilde{\theta}_1, \tilde{\theta}_2 \in \tilde{\Theta}_{j^*}^{m_k}} \int |\tilde{p}(x|\tilde{\theta}_1) - \tilde{p}(x|\tilde{\theta}_2)| d\nu(x) < (1+\epsilon)\alpha. \tag{16}$$

Combining (15) and (16) yields the desired contradiction.

C Computational Details for Applications

C.1 Algorithm

The fixed point iterations described in Section 3 require repeated evaluation of coverage probabilities. These may be computed using an importance sampling approach: Let $\bar{p}$ be a proposal density such that $p(\theta_j|x)$ is absolutely continuous with respect to $\bar{p}$, and let X_i, $i = 1, \ldots, N$ be N i.i.d. draws from $\bar{p}$. Then non-coverage probability of a set φ at θ_j can then be written as $\mathrm{RP}_j = \int \varphi(\theta_j, x) p(\theta_j|x) d\nu(x) = \int \varphi(\theta_j, x) \frac{p(\theta_j|x)}{\bar{p}(x)} \bar{p}(x) d\nu(x)$, yielding the approximation

$$\widehat{\mathrm{RP}}_j(\varphi) = N^{-1} \sum_{i=1}^{N} \varphi(\tilde{\theta}, X_i) \frac{p(\theta_j|X_i)}{\bar{p}(X_i)}.$$

Write φ_π for the set $\varphi(\theta_j, x; \pi)$ of Theorem 1. We employ the following algorithm to obtain an approximate $\pi^\star$ such that the HPD set $\varphi_{\pi^\star}$ has nearly coverage close to the nominal level:

1. Compute and store $\frac{p(\theta_j|X_i)}{\bar{p}(X_i)}$, $i = 1, \ldots, N$, $j = 1, \ldots, m$.

2. Initialize $\pi^{(0)}$ at $\pi_j^{(0)} = 1/m$, $j = 1, \ldots, m$.

3. For $l = 0, 1, \ldots$

   (a) Compute $z_j = \widehat{\mathrm{RP}}_j(\varphi_{\pi^{(l)}}) - \alpha$, $j = 1, \ldots, m$.

   (b) If $\max_j z_j - \min_j z_j < \varepsilon$, set $\pi^\star = \pi^{(l)}$ and end.

   (c) Otherwise, set $\pi_j^{(l+1)} = \exp(\omega z_j) \pi_j^{(l)} / \sum_{k=1}^{m} \exp(\omega z_k) \pi_k^{(l)}$, $j = 1, \ldots, m$, and go to step 3a.

We set $\varepsilon = 0.0003$, and found $\omega = 1.5$ to yield reliable results as long as N is chosen large enough.

In the context of obtaining a credible set with approximately uniform coverage in a bounded but continuous set $\tilde{\Theta}$, we employ the above algorithm for a given partition m with $\varphi_\pi = \varphi^m(\theta_j, x; \pi)$ now defined as described in Section 2.3.3. In addition, we evaluate the uniform coverage properties of the resulting set estimator on $\tilde{\Theta}$, $\tilde{\varphi}^m(\tilde{\theta}, x; \pi^{m\star})$, by computing the (approximate) non-coverage probabilities $\widehat{\mathrm{RP}}(\tilde{\theta}) = N^{-1} \sum_{i=1}^{N} \tilde{\varphi}^m(\tilde{\theta}, X_i; \pi^{m\star}) \frac{p(\tilde{\theta}|X_i)}{\bar{p}(X_i)}$ over a fine grid of values of $\tilde{\theta}$. If these uniform properties are unsatisfactory, then the algorithm is repeated using a finer partition.

For an unbounded parameter space $\tilde{\Theta}$, we implement the guess and verify approach described in Section 2.3.4. We first choose an appropriate κ_S by computing the coverage of the HPD set relative to a flat prior on $\tilde{\Theta}$, and select κ_S to be just large enough for coverage to be sufficiently close to $1-\alpha$ for all $\tilde{\theta}$ with $\varsigma(\tilde{\theta}) > \kappa_S$. We then partition $\tilde{\Theta}_{NS}$ into $m-1$ subsets, and set $\tilde{\Theta}_m^m = \tilde{\Theta}_S$. As discussed in Section 2.3.4, it makes sense to rule out a large discontinuity of $\bar{\Pi}^m(\tilde{\theta}, \pi^{m\star})$ at the boundary between $\tilde{\Theta}_{NS}$ and $\tilde{\Theta}_S$. Thus, in the above algorithm, we directly adjust the $m-1$ values of $\bar{\Pi}^m(\tilde{\theta}, \pi)$ on $\tilde{\Theta}_j^m$, $j = 1, \ldots, m-1$ without any scale normalization, and simply set $\bar{\Pi}^m(\tilde{\theta}, \pi)$ on $\tilde{\Theta}_S$ equal to the value of $\bar{\Pi}^m(\tilde{\theta}, \pi)$ of a subset $\tilde{\Theta}_j^m$ neighboring $\tilde{\Theta}_S$. This has the additional advantage of avoiding computation of the potentially ill-defined RP_m. After the iterations have concluded, we evaluate the coverage properties of the resulting set estimator $\bar{\phi}^m(\tilde{\theta}, x; \pi^{m\star})$ on a fine grid on $\tilde{\Theta}_{NS}$, and on a fine grid in the part of $\tilde{\Theta}_S$ where $\bar{\phi}^m(\tilde{\theta}, x; \pi^{m\star})$ is affected by the shape of the prior

$\bar{\Pi}^m(\tilde{\theta}, \pi)$ on $\tilde{\Theta}_{NS}$. If these are unsatisfactory, we increase m and/or κ_S.

C.2 Details for Break Date and Magnitude

The continuous process X is approximated with 800 steps. Uniform coverage is evaluated on the Cartesian grid with $\lambda \in \{0.15, 0.15125, 0.1525, \ldots, 0.85\}$ and $\delta \in \{-15.0, -14.99, -14.98, \ldots, 15\}$. We set $N = 3 \cdot 10^6$, and $\bar{p}$ to be uniform on $\{(\lambda, \delta) : 0.13 \leq \lambda \leq 0.87, \ -16 \leq \delta \leq 16\}$, which yields Monte Carlo standard deviations of coverage probabilities uniformly smaller than 0.001. We impose symmetry with respect to the sign of δ, and around $\lambda = 0.5$, in the computation of $\pi^{m\star}$. Computations for the finer partition take about 6 hours on a modern PC.

In the application, we use Elliott and Müller's (2014) estimate of 2.6 for the long-run standard deviation of the quarterly data y_t.

C.3 Details for Autoregressive Root Near Unity

The continuous process X is approximated with 800 steps. Uniform coverage is evaluated on 5001 values $\tilde{\theta} \in \{120(\frac{j}{5000})^2\}_{j=0}^{5000}$. We set $N = 1.5 \cdot 10^6$, and set $\bar{p}$ to be uniform on the 101 values $\tilde{\theta} \in \{160(\frac{j}{100})^2\}_{j=0}^{100}$, which yields Monte Carlo standard deviations of coverage probabilities uniformly smaller than 0.001.

For the application, we rely on output of the DF-GLS regression also employed in Lopez, Murray, and Papell (2013) to obtain small sample analogues to $\int_0^1 X(s)dX(s)$ and $\int_0^1 X(s)^2 ds$. Specifically, let $\hat{\rho}$, $\hat{\sigma}_\rho$ and $\hat{\phi}$ be the usual OLS estimate of ρ, its standard error and the additional coefficients in an augmented Dickey-Fuller regression using GLS demeaned data (with lag length as determined by Lopez, Murray, and Papell (2013)). We then employ the analogue $(T^{-1}\hat{\phi}(1)(\hat{\rho}-1)/\hat{\sigma}^2_{\rho}, T^{-2}\hat{\phi}(1)^2/\hat{\sigma}^2_{\rho}) \Rightarrow (\int_0^1 X(s)dX(s), \int_0^1 X(s)^2 ds)$ for the empirical results in Table 1.