Time Varying Structural Vector Autoregressions and Monetary Policy: Appendix to the Corrigendum

Marco Del Negro Giorgio E. Primiceri

February 2015

In this appendix, we (i) re-estimate the model of Primiceri (2005) using Algorithm 2 (the sampler from the approximate posterior) and Algorithm 3 (the sampler from the true posterior), and compare these results with those obtained with Algorithm 1 (the original algorithm of Primiceri, 2005); (ii) present a more formal treatment of Algorithm 2 and Algorithm 3; (iii) formally explain why Algorithm 1 provides a poor approximation of the posterior distribution; and (iv) apply Geweke's (2004) "Joint Distribution Tests of Posterior Simulators" to Algorithm 1, 2 and 3, and present the results of these tests.

1 New Results Based on Algorithm 2 and 3

In this section, we reproduce the figures of Primiceri (2005) using Algorithm 2 (the sampler from the approximate posterior) and Algorithm 3 (the sampler from the true posterior), and compare these results with those obtained with Algorithm 1 (the original algorithm of Primiceri, 2005). The new results are based on 70,000 draws of the Gibbs sampler, discarding the first 20,000 to allow for convergence to the ergodic distribution.

The first thing to notice is that the results based on Algorithm 3 are qualitatively similar to the original ones obtained with Algorithm 1, but they are not the same (figures 1-8). The main difference is that some estimates of the time-varying objects are now smoother. For

example, the standard deviation of monetary policy shocks (figure 1c) exhibits substantial time variation, but not as much as in the original results. A similar comment applies to the interest rate response to a permanent increase in inflation and unemployment (figures 5 and 7).

The results obtained using Algorithm 2 and 3 are instead indistinguishable from each other (figures 9-16), suggesting that the mixture-of-normals approximation error involved in the procedure of Kim, Shephard and Chib (1998, KSC hereafter) is negligible in our application (as it was in theirs).

2 A Formal Treatment of Algorithm 2 and 3

In this section, we present a formal derivation of Algorithm 2 and 3.

2.1 Algorithm 2

The joint posterior distribution of Σ^T and θ is given by

$$p(\Sigma^T, \theta | y^T) \propto p(y^T | \Sigma^T, \theta) \cdot p(\Sigma^T, \theta) \tag{1}$$

where $p(y^T|\Sigma^T, \theta)$ is the likelihood function implied by equation (1.1) of the corrigendum, and $p(\Sigma^T, \theta)$ is the prior density of Σ^T and θ. In principle, one could use a two-block Gibbs sampler in Σ^T and θ with steps: i) draw Σ^T from $p(\Sigma^T|y^T, \theta) \propto p(y^T|\Sigma^T, \theta) \cdot p(\Sigma^T|\theta)$, and ii) draw θ from $p(\theta|y^T, \Sigma^T) \propto p(y^T|\Sigma^T, \theta) \cdot p(\theta|\Sigma^T)$. While step (ii) is straightforward, step (i) is not: the time-varying volatilities Σ^T enter the model multiplicatively, making it impossible to use linear and Gaussian state-space methods. KSC's idea consists of approximating the likelihood $p(y^T|\theta, \Sigma^T)$ using the mixture-of-normals $\int \tilde{p}(y^T|\Sigma^T, \theta, s^T)\pi(s^T)ds^T$, where s^T represents the components of the mixture for each date and variable, $\pi(s^T)$ are the corresponding mixture weights, and $\tilde{p}(y^T|\Sigma^T, \theta, s^T)$ is the likelihood of the data conditional on the mixture components s^T.[1]

[1]Note that there is a clear abuse of notation in writing $\int \tilde{p}(y^T|\theta, \Sigma^T, s^T)\pi(s^T)ds^T$ given that the s^T are discrete indicators, but this shortcut simplifies the notation considerably.

Let $\tilde{p}(\Sigma^T, \theta, s^T|y^T)$ denote the product of $\tilde{p}(y^T|\Sigma^T, \theta, s^T)$ and the prior of θ, Σ^T and s^T, that is

$$\tilde{p}(\Sigma^T, \theta, s^T|y^T) = \tilde{p}(y^T|\Sigma^T, \theta, s^T) \cdot p(\Sigma^T, \theta) \cdot \pi(s^T). \tag{2}$$

In addition, for the sake of argument, suppose that the mixture-of-normals provides a perfect approximation of the likelihood, i.e. that $p(y^T|\theta, \Sigma^T) = \int \tilde{p}(y^T|\Sigma^T, \theta, s^T)\pi(s^T)ds^T$. Integrating out the mixture components s^T from $\tilde{p}(\Sigma^T, \theta, s^T|y^T)$ in (2), we obtain $p(y^T|\theta, \Sigma^T) \cdot p(\Sigma^T, \theta)$, which is proportional to the the posterior of interest $p(\theta, \Sigma^T|y^T)$. This implies that, if we device an algorithm for drawing from $\tilde{p}(\Sigma^T, \theta, s^T|y^T)$, after discarding the draws of s^T, we are left with draws of θ and Σ^T from the desired distribution. *Algorithm 2*, which we rewrite below, represents such an algorithm:

1. Draw Σ^T from $\tilde{p}\left(\Sigma^T|y^T, \theta, s^T\right) \propto \tilde{p}(y^T|\Sigma^T, \theta, s^T) \cdot p(\Sigma^T|\theta)$
2. Draw (θ, s^T) from $\tilde{p}\left(\theta, s^T|y^T, \Sigma^T\right)$, which is accomplished by
   (a) Drawing θ from $p\left(\theta|y^T, \Sigma^T\right) \propto p(y^T|\Sigma^T, \theta) \cdot p(\theta|\Sigma^T)$.
   (b) Drawing s^T from $\tilde{p}\left(s^T|y^T, \Sigma^T, \theta\right) \propto \tilde{p}(y^T|\Sigma^T, \theta, s^T) \cdot \pi(s^T)$.

As emphasized in the note, Algorithm 2 is conceived as a two-blocks sampler, with blocks Σ^T and (θ, s^T). We draw from the joint of (θ, s^T) given Σ^T and y^T by first drawing from the marginal $p\left(\theta|y^T, \Sigma^T\right)$ and then from the conditional $\tilde{p}\left(s^T|y^T, \Sigma^T, \theta\right)$. It is precisely the fact that we draw from the marginal of θ that allows us to use the original likelihood $p(y^T|\Sigma^T, \theta)$ in step 2a: under the assumption that there is no approximation error, integrating out the s^T from the joint distribution (2) yields

$$p(\Sigma^T, \theta) \cdot \int \tilde{p}(y^T|\Sigma^T, \theta, s^T)\pi(s^T)ds^T = p(y^T|\Sigma^T, \theta) \cdot p(\Sigma^T, \theta) \propto p(y^T|\Sigma^T, \theta) \cdot p(\theta|\Sigma^T).$$

Furthermore, step 1 is also simple: as discussed in the paper, conditional on s^T, the model is linear and Gaussian in the log-volatilities, making the distribution $\tilde{p}(y^T|\Sigma^T, \theta, s^T)$ amenable to the use of linear and Gaussian state-space methods.

2.2 Algorithm 3

In the previous section we have provided a justification for Algorithm 2 under the assumption that $p(y^T|\theta, \Sigma^T) = \int \tilde{p}(y^T|\Sigma^T, \theta, s^T)\pi(s^T)ds^T$. Of course, in practice, this is not correct: the mixture of normals is only an approximation of the true likelihood. In this subsection we present a formal treatment of Algorithm 3, which addresses this issue.

Construct a joint posterior of Σ^T, θ and s^T as follows:

$$\begin{aligned} p(\theta, \Sigma^T, s^T|y^T) &= p(\theta, \Sigma^T|y^T) \cdot \tilde{p}(s^T|\Sigma^T, \theta, y^T) \\ &\propto p(y^T|\theta, \Sigma^T) \cdot p(\Sigma^T, \theta) \cdot \tilde{p}(s^T|\Sigma^T, \theta, y^T), \end{aligned} \tag{3}$$

with

$$\tilde{p}(s^T|\Sigma^T, \theta, y^T) = \frac{\tilde{p}(y^T|\Sigma^T, \theta, s^T) \cdot \pi(s^T)}{c(\Sigma^T, \theta, y^T)}, \tag{4}$$

where $c(\Sigma^T, \theta, y^T) \equiv \int \tilde{p}(y^T|\Sigma^T, \theta, s^T)\pi(s^T)ds^T$ guarantees that the density in (4) integrates to one.

As discussed above, a perfectly fine approach for obtaining draws from the posterior of interest, $p(\theta, \Sigma^T|y^T)$, is to sample from $p(\theta, \Sigma^T, s^T|y^T)$, and then discard the draws of s^T. This is precisely what Algorithm 3 does. Like Algorithm 2, Algorithm 3 has the structure of a two-block sampler, with blocks Σ^T and $\left(\theta, s^T\right)$. However, Algorithm 3 follows Stroud et al. (2003) in using a Metropolis-Hastings step for drawing Σ^T conditional on (θ, s^T), where the proposal pdf is given by

$$\tilde{p}\left(\Sigma^T|y^T, \theta, s^T\right) \propto \tilde{p}(y^T|\Sigma^T, \theta, s^T) \cdot p(\Sigma^T|\theta), \tag{5}$$

which is the density used in step 1 of Algorithm 2.[2]

Specifically, *Algorithm 3* consists of the following steps:

[2]Stroud et al. (2003) study the use of mixture approximations in Gibbs samplers, and thus generalize the results of KSC.

1. Draw Σ^T from $p\left(\Sigma^T|y^T,\theta,s^T\right)$ as follows: Draw a candidate $\tilde{\Sigma}^T$ from the proposal density $\tilde{p}\left(\Sigma^T|y^T,\theta,s^T\right)$ of Algorithm 2, and set

$$\Sigma^{(j)\ T}=\begin{cases}\tilde{\Sigma}^T & \text{with probability } \alpha\\ \Sigma^{(j-1)\ T} & \text{with probability } 1-\alpha\end{cases},$$

where the superscript (j) denotes the iteration of the sampler, and where

$$\alpha=\frac{p\left(\tilde{\Sigma}^T|y^T,\theta,s^T\right)}{p\left(\Sigma^{(j-1)\ T}|y^T,\theta,s^T\right)}\frac{\tilde{p}\left(\Sigma^{(j-1)\ T}|y^T,\theta,s^T\right)}{\tilde{p}\left(\tilde{\Sigma}^T|y^T,\theta,s^T\right)}.$$

2. Draw (θ,s^T) from $p\left(\theta,s^T|y^T,\Sigma^T\right)$, which is accomplished by

   (a) Drawing θ from

$$\begin{aligned}p\left(\theta|y^T,\Sigma^T\right) &= \int p(\theta,s^T|y^T,\Sigma^T)ds^T\\ &\propto p(y^T|\theta,\Sigma^T)\cdot p(\theta|\Sigma^T)\cdot\int\tilde{p}(s^T|\Sigma^T,\theta,y^T)ds^T=p(y^T|\Sigma^T,\theta)\cdot p(\theta|\Sigma^T).\end{aligned}$$

   (b) Drawing s^T from $\tilde{p}\left(s^T|y^T,\Sigma^T,\theta\right)\propto\tilde{p}(y^T|\Sigma^T,\theta,s^T)\cdot\pi(s^T)$.

Observe that, since step 1 takes θ and s^T as given, the acceptance probability can be rewritten as

$$\alpha=\frac{p\left(\tilde{\Sigma}^T,\theta,s^T|y^T\right)}{p\left(\Sigma^{(j-1)\ T},\theta,s^T|y^T\right)}\frac{\tilde{p}\left(\Sigma^{(j-1)\ T}|y^T,\theta,s^T\right)}{\tilde{p}\left(\tilde{\Sigma}^T|y^T,\theta,s^T\right)}.$$

Using (3), (4) and (5), we then obtain

$$\alpha=\frac{p(y^T|\theta,\tilde{\Sigma}^T)}{p(y^T|\theta,\Sigma^{(j-1)\ T})}\frac{c(\Sigma^{(j-1)\ T},\theta,y^T)}{c(\tilde{\Sigma}^T,\theta,y^T)}.$$

Finally, notice that $c(\Sigma^T,\theta,y^T)$ coincides with the mixture-of-normals approximation of the original likelihood $p(y^T|\Sigma^T,\theta)$, hence

$$\alpha=\frac{\left(\prod_t\phi\left(y_t^*|0_{n\times1},\tilde{\Sigma}_t\tilde{\Sigma}_t'\right)\right)\left(\prod_t\prod_i mn_{KSC}\left(y_{i,t}^{**}-2\log\sigma_{i,t}^{(j-1)}\right)\right)}{\left(\prod_t\phi\left(y_t^*|0_{n\times1},\Sigma_t^{(j-1)}\Sigma_t^{(j-1)\ \prime}\right)\right)\left(\prod_t\prod_i mn_{KSC}\left(y_{i,t}^{**}-2\log\tilde{\sigma}_{i,t}\right)\right)},\tag{6}$$

where $y_t^*=A_t\left(y_t-c_t-B_{1,t}y_{t-1}-...-B_{k,t}y_{t-k}\right)$, $y_{i,t}^{**}=\log\left(y_{i,t}^{*2}+0.001\right)$, $\sigma_{i,t}$ is the i-th element of the diagonal of Σ_t, $\phi\left(\cdot|0_{n\times1},\tilde{\Sigma}_t\tilde{\Sigma}_t'\right)$ is the pdf of an $n-$variate Gaussian distribution with mean zero and variance $\tilde{\Sigma}_t\tilde{\Sigma}_t'$, and $mn_{KSC}\left(\cdot\right)$ denotes the pdf of the mixture-of-normals distribution with means, variances and mixing proportions specified in KSC.

3 The Fixed-Point Integral Equation

In this section, we formally explain why the original algorithm of Primiceri (2005) is not a proper Gibbs sampling. The reason can be understood from inspecting the key equation showing why the Markov chain converges (we omit the conditioning on y^T to simplify notation):

$$p(\theta, \Sigma^T) = \int h\left(\theta, \Sigma^T | \theta', \Sigma^{T\prime}\right) p\left(\theta', \Sigma^{T\prime}\right) \mathrm{d}\left(\theta', \Sigma^{T\prime}\right) \tag{7}$$

where $h\left(\theta, \Sigma^T | \theta', \Sigma^{T\prime}\right)$ is a Markov transition kernel defined by

$$h\left(\theta, \Sigma^T | \theta', \Sigma^{T\prime}\right) = \int p(\theta, \Sigma^T | s^T) p(s^T | \theta', \Sigma^{T\prime}) \mathrm{d}s^T. \tag{8}$$

Equation (7) defines a fixed point integral equation for which the true marginal $p(\theta, \Sigma^T)$ is a solution, which is readily seen by plugging (8) into (7) and changing the order of integration, as shown below (Chib and Greenberg, 1996 and references therein discuss why i) it is the unique solution, and ii) there is convergence from any initial $p\left(\theta', \Sigma^{T\prime}\right)$ under general conditions). Omitting to condition on s^T when drawing from $p(\theta, \Sigma^T | s^T)$ (as done in step 3) implies using the wrong kernel, hence the fixed point argument breaks down: even if one were to draw $(\theta', \Sigma^{T\prime})$ from the correct joint distribution, the resulting (θ, Σ^T) in the next iteration would not be from $p(\theta, \Sigma^T)$.

In the three-block Gibbs sampler, equation (3.1)—the fixed point integral equation—becomes

$$p(\theta, \Sigma^T, s^T) = \int .. \int h\left(\theta, \Sigma^T, s^T | \theta', \Sigma^{T'}, s^{T'}\right) p\left(\theta', \Sigma^{T'}, s^{T'}\right) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'} \tag{9}$$

where $h\left(\theta, \Sigma^T, s^T | \theta', \Sigma^{T'}, s^{T'}\right)$ is a Markov transition kernel defined by

$$h\left(\theta, \Sigma^T, s^T | \theta', \Sigma^{T'}, s^{T'}\right) = p(\theta | \Sigma^T, s^T) p(\Sigma^T | \theta', s^T) p(s^T | \theta', \Sigma^{T'}). \tag{10}$$

Here we follow Chib and Greenberg (1996) and show that $p(\theta, \Sigma^T, s^T)$ is indeed the solution to (9). In fact, one can write the right hand side of expression (9), after substituting in the definition of the transition kernel (10), as:

$$\int .. \int p(\theta | \Sigma^T, s^T) p(\Sigma^T | \theta', s^T) p(s^T | \theta', \Sigma^{T'}) p\left(\theta', \Sigma^{T'}, s^{T'}\right) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'} =$$
$$\int .. \int p(\theta | \Sigma^T, s^T) \frac{p(\Sigma^T | s^T) p(\theta' | \Sigma^T, s^T)}{p(\theta' | s^T)} \frac{p(s^T) p(\theta', \Sigma^{T'} | s^T)}{p(\theta', \Sigma^{T'})} p\left(\theta', \Sigma^{T'}, s^{T'}\right) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'}$$

where we used Bayes law to express $p(\Sigma^T|\theta', s^T)$ and $p(s^T|\theta', \Sigma^{T'})$. Note that the terms

$$p(\theta|\Sigma^T, s^T)p(\Sigma^T|s^T)p(s^T) = p(\theta, \Sigma^T, s^T)$$

can be taken out of the integral as they do not depend on the $'$ variables, and their product is precisely $p(\theta, \Sigma^T, s^T)$. Therefore we just have to show that

$$\int .. \int \frac{p(\theta'|\Sigma^T, s^T)}{p(\theta'|s^T)} \frac{p(\theta', \Sigma^{T'}|s^T)}{p(\theta', \Sigma^{T'})} p\left(\theta', \Sigma^{T'}, s^{T'}\right) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'} = 1.$$

This is the case because

$$\begin{aligned}
&\int .. \int \frac{p(\theta'|\Sigma^T, s^T)}{p(\theta'|s^T)} \frac{p(\theta', \Sigma^{T'}|s^T)}{p(\theta', \Sigma^{T'})} p(\theta', \Sigma^{T'}, s^{T'}) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'} = \\
&\quad \int .. \int \frac{p(\theta'|\Sigma^T, s^T)}{p(\theta'|s^T)} \frac{p(\theta'|s^T)p(\Sigma^{T'}|\theta', s^T)}{p(\theta', \Sigma^{T'})} p(\theta', \Sigma^{T'}) p(s^{T'}|\theta', \Sigma^{T'}) \mathrm{d}\theta' \mathrm{d}\Sigma^{T'} \mathrm{d}s^{T'} = \\
&\qquad \int p(\theta'|\Sigma^T, s^T) \left(\int p(\Sigma^{T'}|\theta', s^T) \left(\int p(s^{T'}|\Sigma^{T'}, s^{T'}) \mathrm{d}s^{T'} \right) \mathrm{d}\Sigma^{T'} \right) \mathrm{d}\theta' = 1,
\end{aligned}$$

where in the second line we again used Bayes law and in the fourth line we realized that we are left with three conditional distributions, all integrating to one. Clearly, omitting to condition on s^T when drawing from $p\left(\theta|\Sigma^T, s^T\right)$ implies using the wrong kernel, and the fixed-point arguments breaks down.

4 Geweke's (2004) "Getting It Right"

In this section, we apply Geweke's (2004) "Joint Distribution Tests of Posterior Simulators" to the three algorithms discussed in the note, and present further evidence that Algorithm 3 is fully correct, Algorithm 2 provides a close approximation to the true posterior distribution, while Algorithm 1 provides a poor approximation. Geweke's idea is to compare two ways of obtaining draws from the joint distribution of the data and the model parameters, $p\left(y^T, \theta, \Sigma^T, s^T\right)$:

a. Draw the parameters from the prior, and then the data from the data-generating process (that is, draw sequentially from $p\left(\theta, \Sigma^T, s^T\right)$ and $p\left(y^T|\theta, \Sigma^T, s^T\right)$).

b. Draw from the posterior using the MCMC algorithm given a draw of the data, and then use this draw to generate another draw from data, and so on (that is, draw sequentially from $p\left(\theta, \Sigma^T, s^T | y^T\right)$ and then from $p\left(y^T | \theta, \Sigma^T, s^T\right)$).

If the MCMC algorithm is correct, (a) and (b) should yield the same distribution, and in particular the same marginal for the model parameters (which, in the case of (a), is of course the prior). Therefore, if the MCMC algorithm is correct, P-P plots constructed using the draws from (a) and (b) should lie on the 45-degree line.

We now present the results obtained by applying this procedure to the various algorithms that we have discussed so far. Note that, for computational reasons, we use $T = 10$ in running these tests, which is smaller than the actual sample size. For a T as large as that in the sample, it simply takes so many draws for (b) to converge (even if the MCMC algorithm is right) that the test is computationally not feasible. Since Geweke's approach applies to any T, we are justified in using a smaller T that makes the comparison feasible.

We concentrate on the P-P plots for the distribution of the log-volatilities at a particular point in time ($t = 7$), because the differences are smaller for the other coefficients. Figure 17 shows the results related to the original algorithm (Algorithm 1). It is evident that the P-P plots are very far from the 45-degree line, indicating that the draws generated with (a) and (b) belong to different distributions. This suggests that Algorithm 1 generates draws from a distribution which is quite different from the true posterior, as we have argued above.

Figure 18 plots the results obtained using Algorithm 2. The fact that the P-P plots in figure 18 are now much closer to the 45-degree line is a sign of dramatic improvement in the accuracy of the algorithm. The natural question is of course why these P-P plots do not lie exactly on top of the 45-degree line, but just close to it. This is due to the minor error involved in the mixture-of-normals approximation proposed by KSC. A property of the Geweke (2004) approach is that it amplifies subtle discrepancies in the sampler, such as these small approximation errors. Figure 19 confirms this conjecture by presenting the P-P plots obtained by running the Geweke procedure using Algorithm 3. In this case, the P-P plots essentially coincide with the 45-degree lines, which verifies that there is no problem

with Algorithm 2, other than the fact that it uses the mixture approximation to increase efficiency and speed of convergence. Recall from section 1 that this approximation is absolutely inconsequential for the estimation results, i.e. for the construction the posterior distribution given the observed data. Conversely, applying the same correction for the mixture-of-normals approximation error in step 1 of the original algorithm does not improve the P-P plots at all, as shown in figure 20.

REFERENCES

CHIB, S. and GREENBERG, E. (1996), "Markov Chain Monte Carlo Simulation Methods in Econometrics," Econometric Theory, 12, 409-431.

GEWEKE, J. (2004), "Getting It Right: Joint Distribution Tests of Posterior Simulators," Journal of the American Statistical Association, 99, 799-804.

KIM, S., SHEPHARD, N. and CHIB, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.

PRIMICERI, G. E. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72(3), 821–852.

STROUD, J.R., MÜLLER, P. and POLSON, N.G. (2003), "Nonlinear State-Space Models With State-Dependent Variances," *Journal of the American Statistical Association*, 98, 377-386.

Figure 1: Posterior mean, 16th and 84th percentiles of the standard deviation of (a) the residuals of the inflation equation, (b) the residuals of the unemployment equation and (c) the residuals of the interest rate equation or monetary policy shocks.

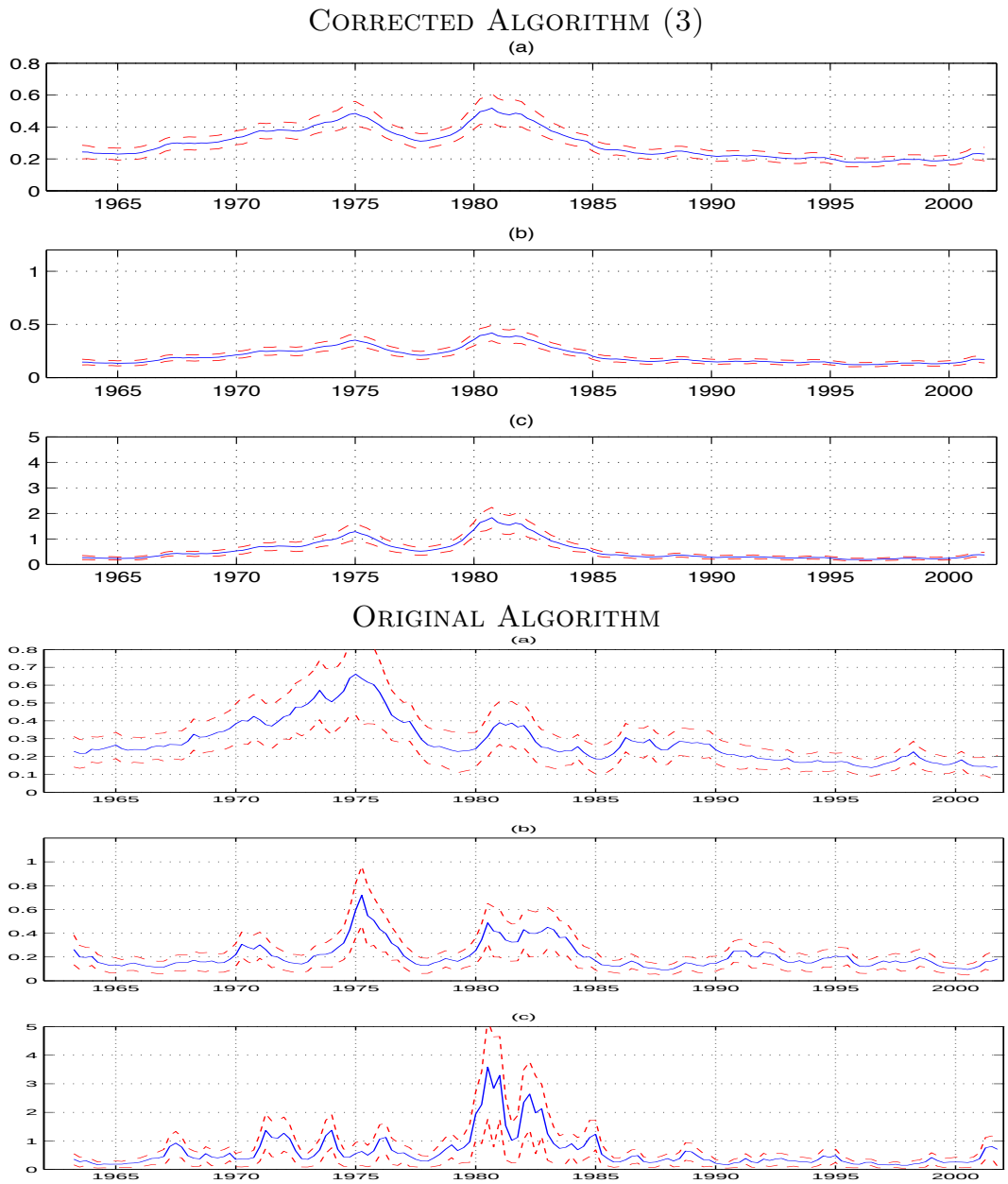

Figure 2: (a) impulse responses of inflation to monetary policy shocks in 1975:I, 1981:III and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

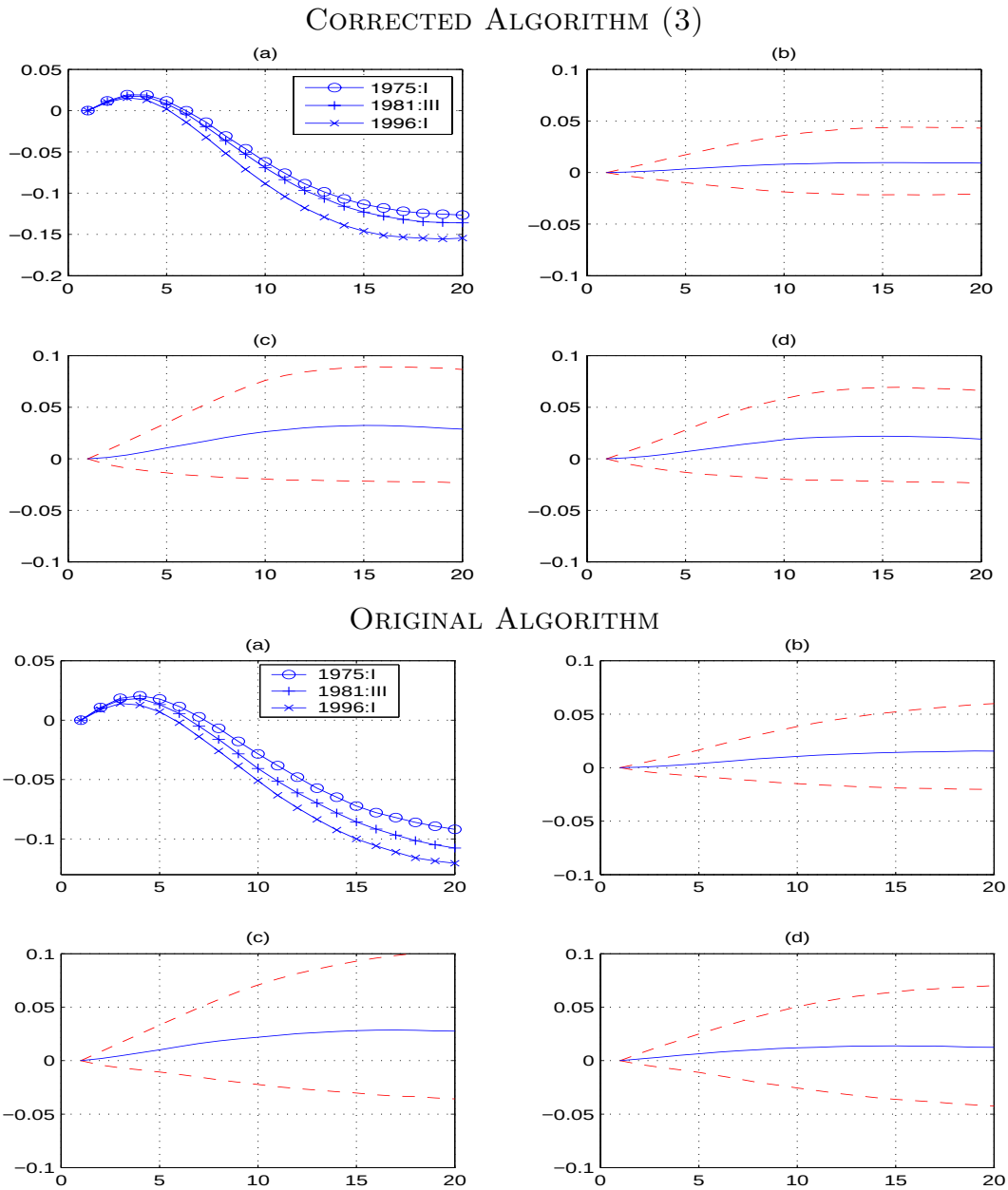

Figure 3: (a) impulse responses of unemployment to monetary policy shocks in 1975:I, 1981:III and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

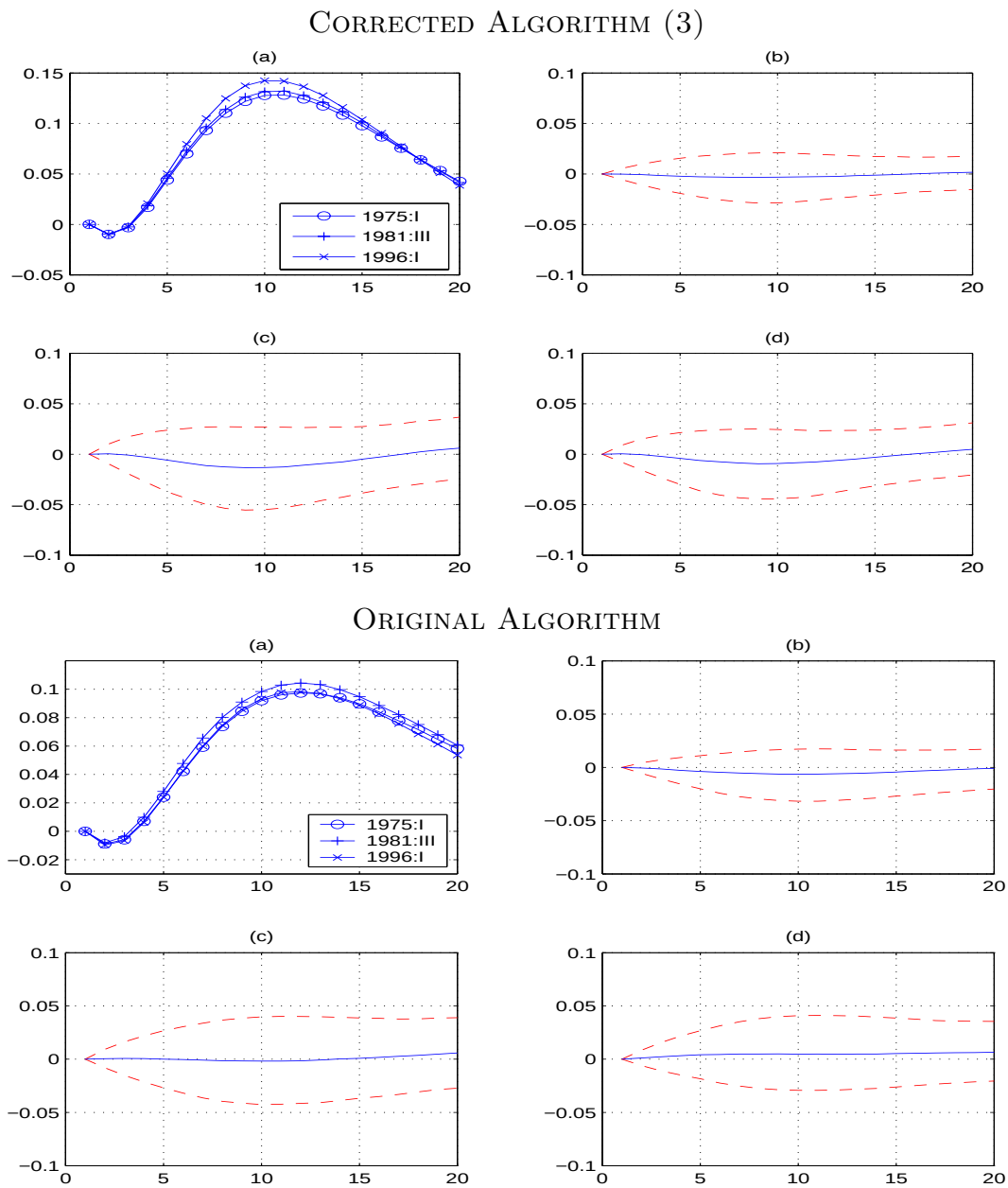

Figure 4: Interest rate response to a 1% permanent increase of inflation with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

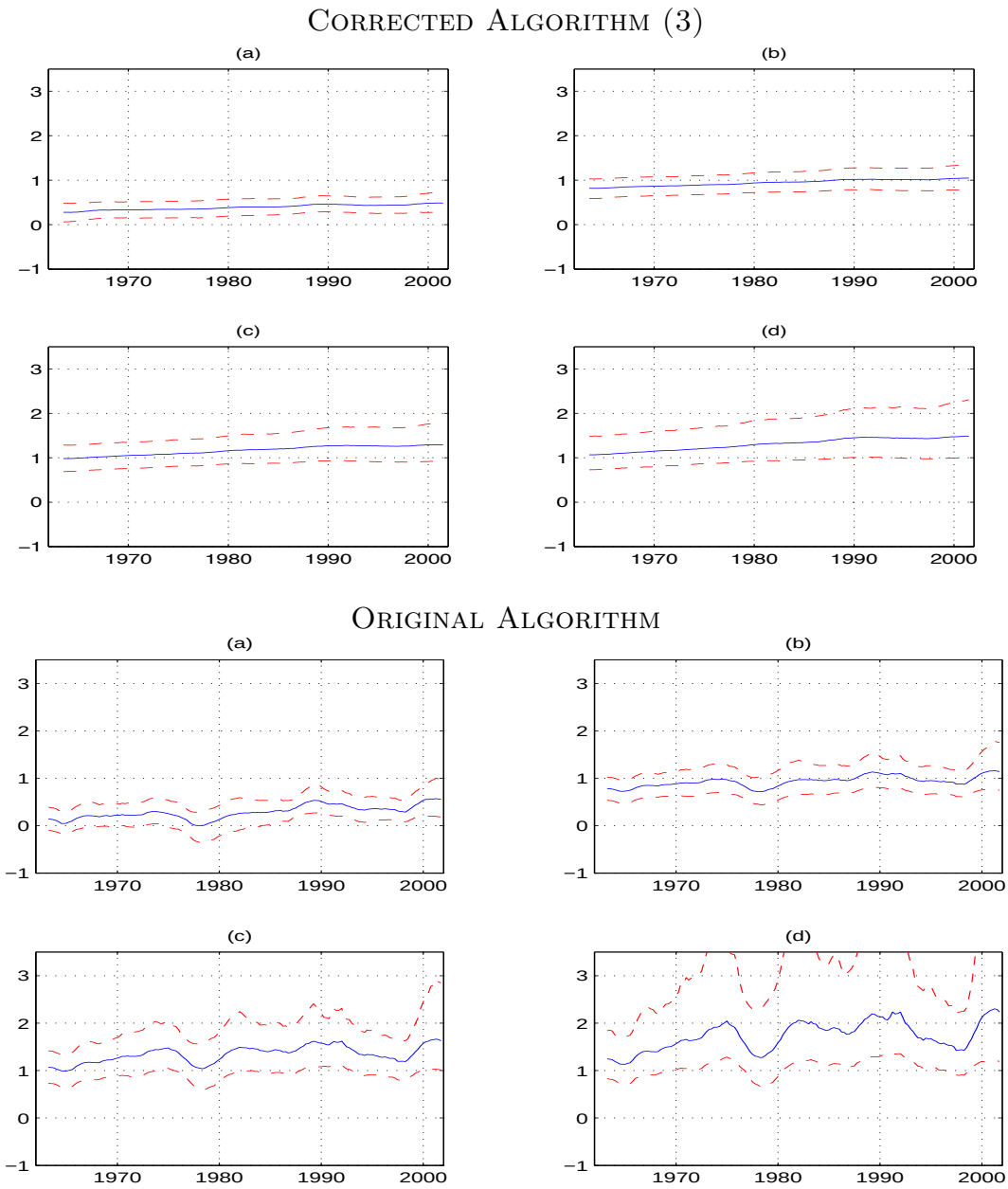

Figure 5: Interest rate response to a 1% permanent increase of unemployment with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

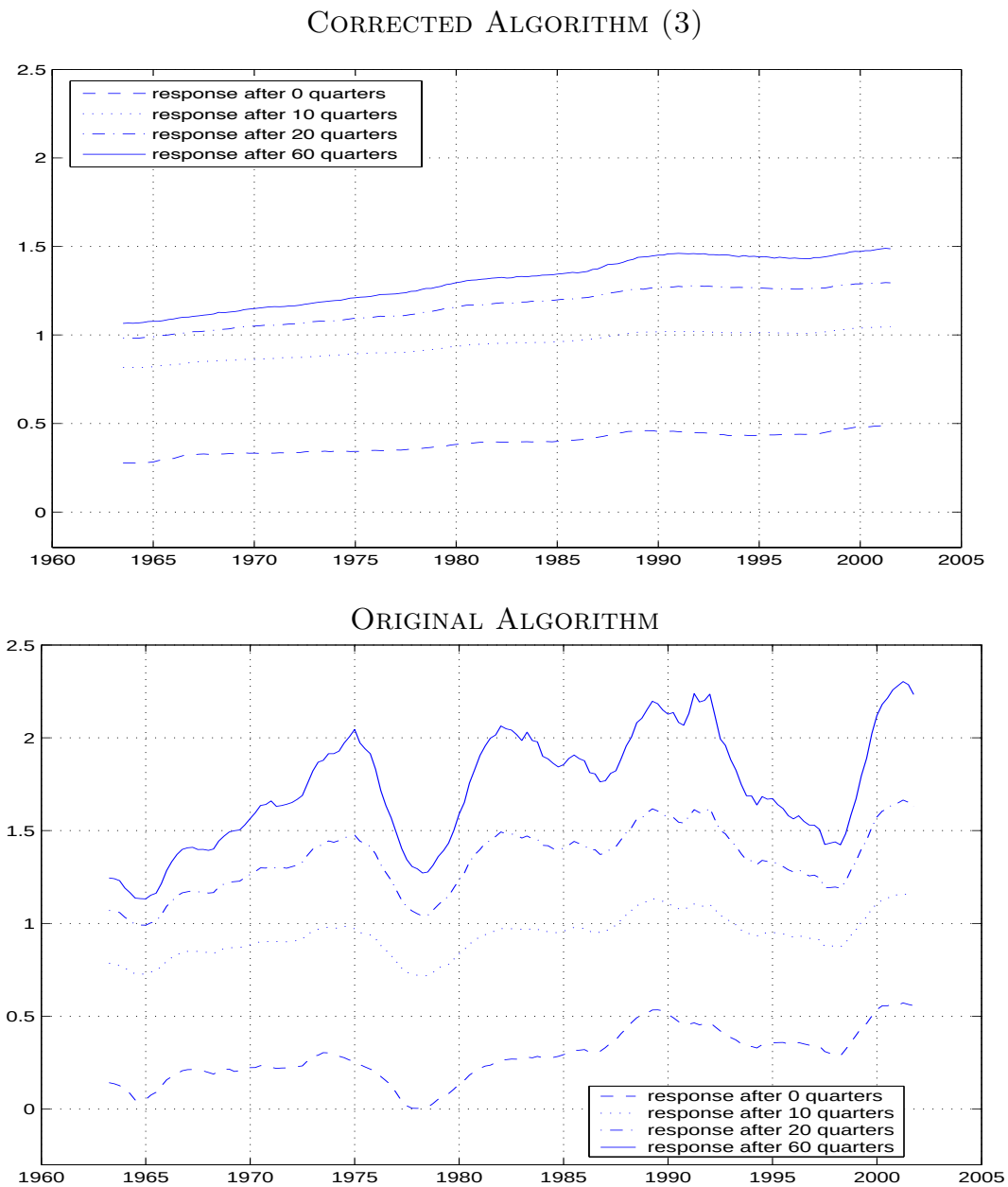

Figure 6: Interest rate response to a 1% permanent increase of unemployment with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

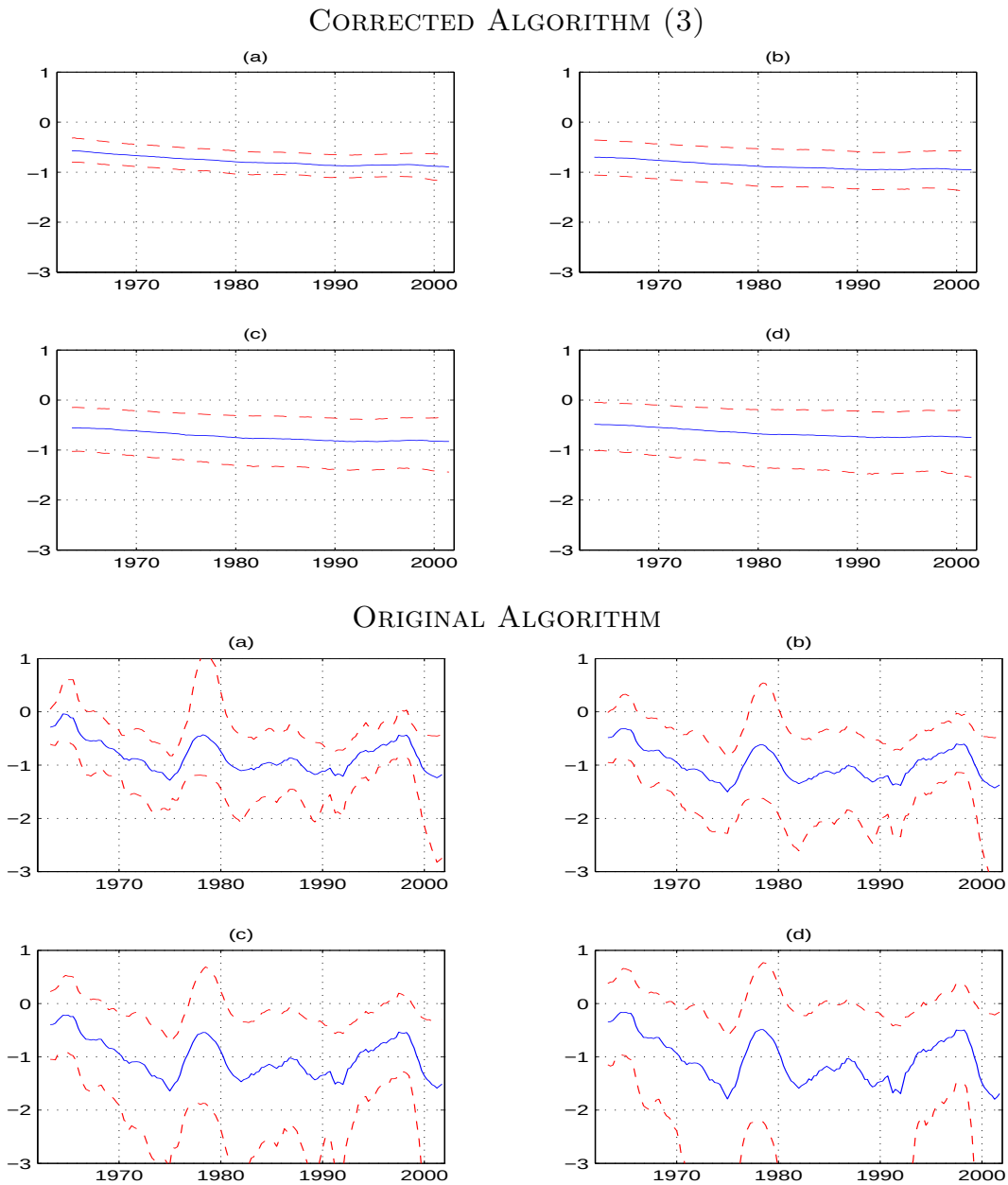

Figure 7: Interest rate response to a 1% permanent increase of unemployment.

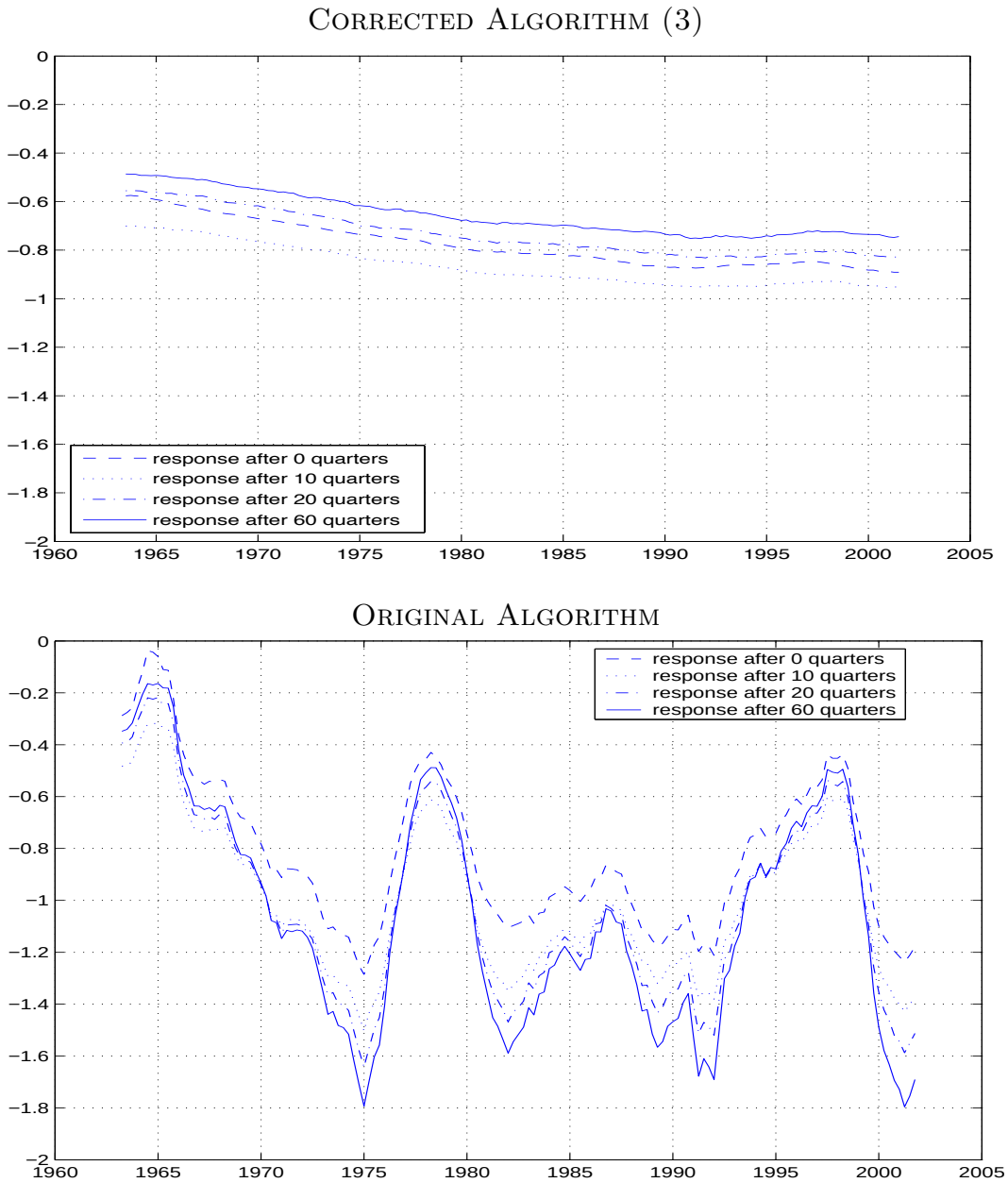

Figure 8: Counterfactual historical simulation drawing the parameters of the monetary policy rule from their 1991-1992 posterior. (a) Inflation, (b) unemployment.

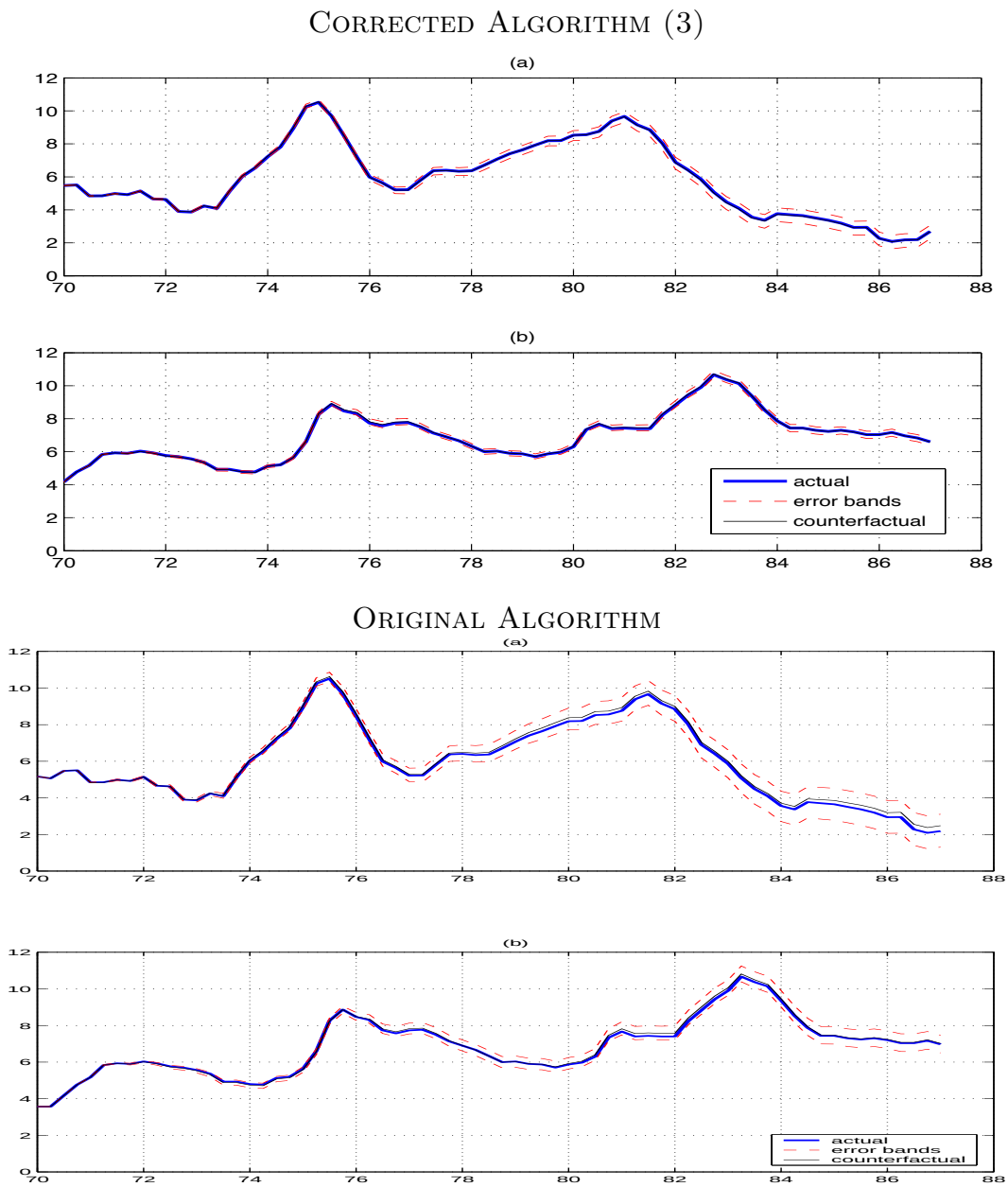

Figure 9: Posterior mean, 16th and 84th percentiles of the standard deviation of (a) the residuals of the inflation equation, (b) the residuals of the unemployment equation and (c) the residuals of the interest rate equation or monetary policy shocks.

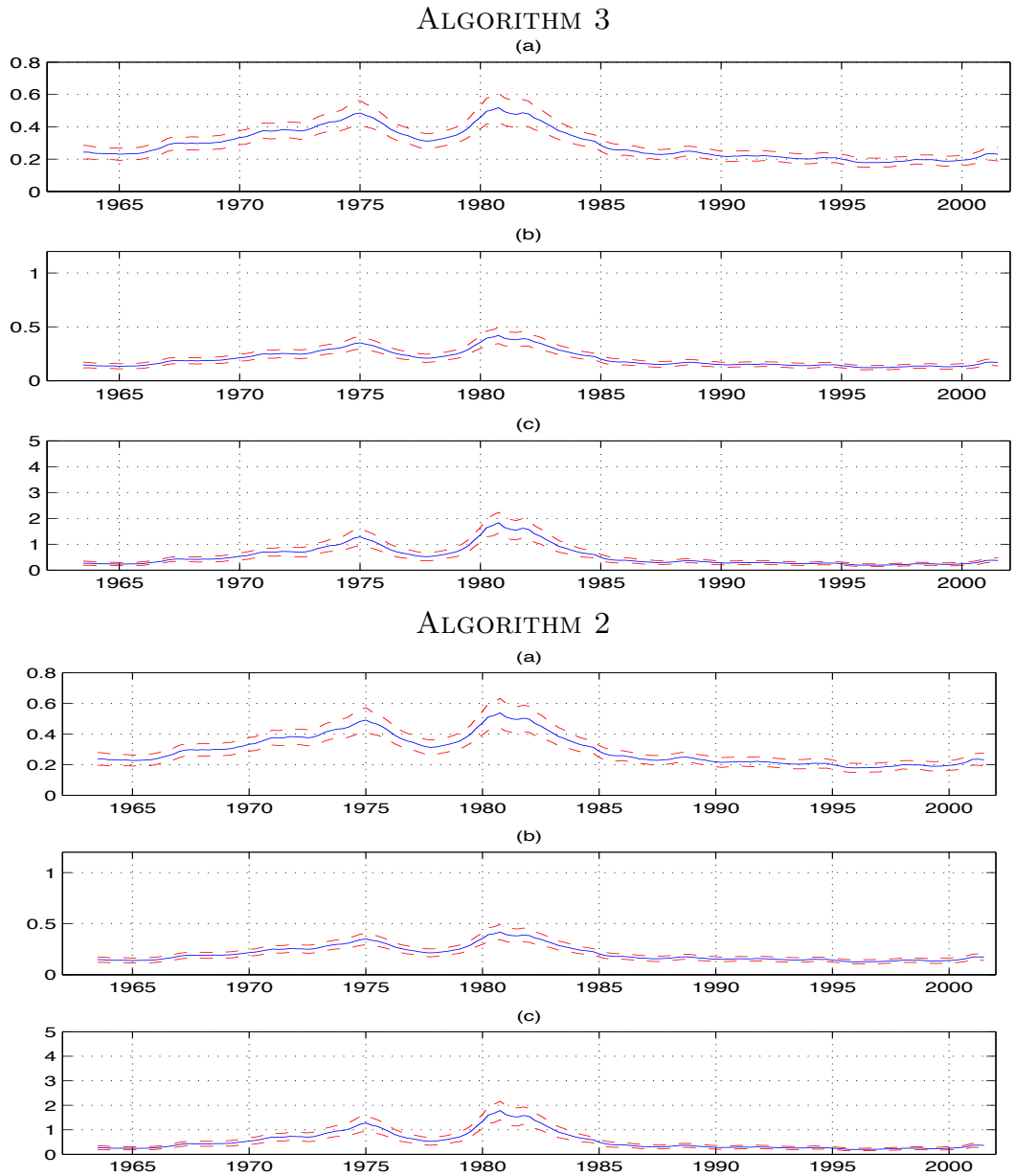

Figure 10: (a) impulse responses of inflation to monetary policy shocks in 1975:I, 1981:III and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

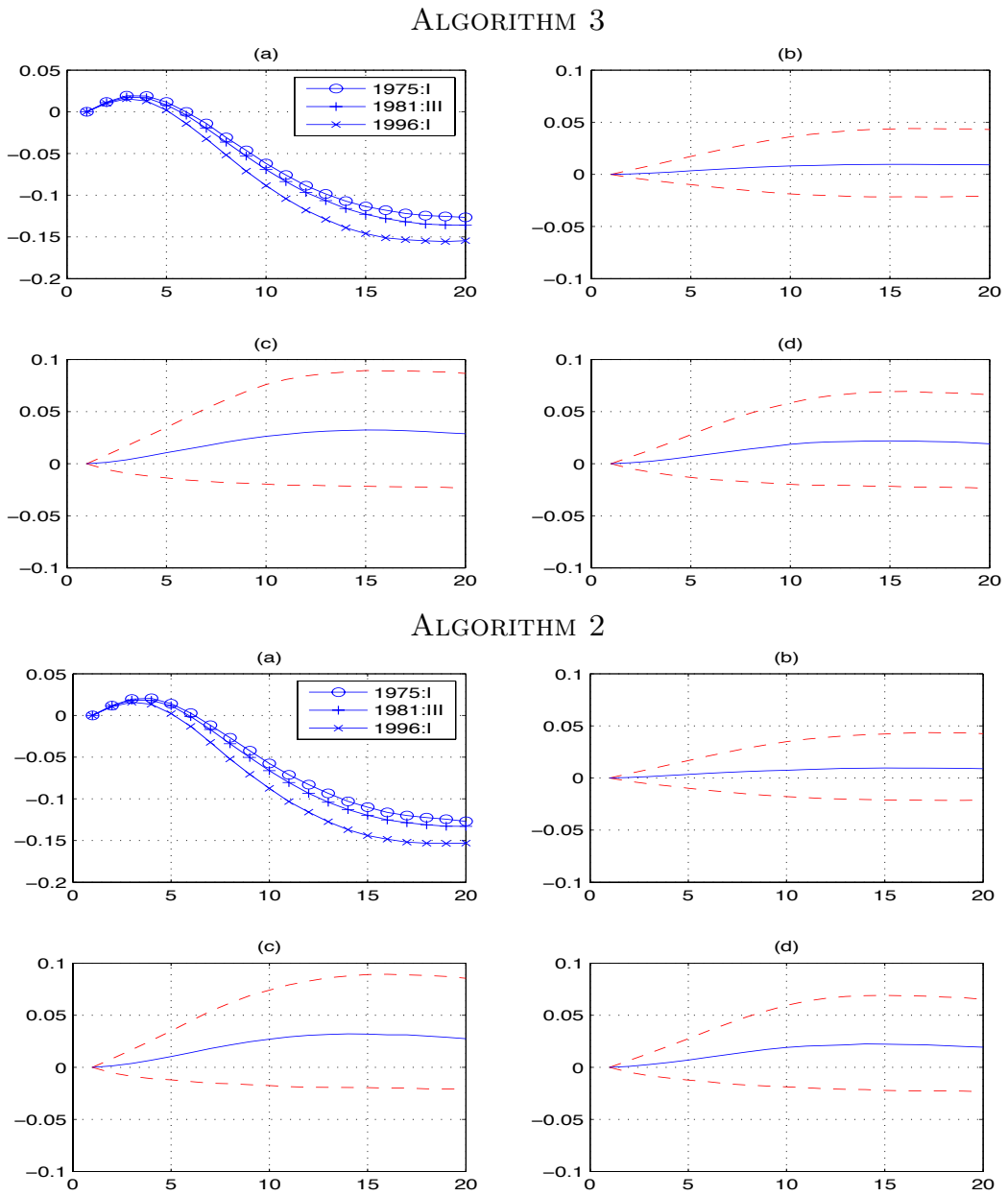

Figure 11: (a) impulse responses of unemployment to monetary policy shocks in 1975:I, 1981:III and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

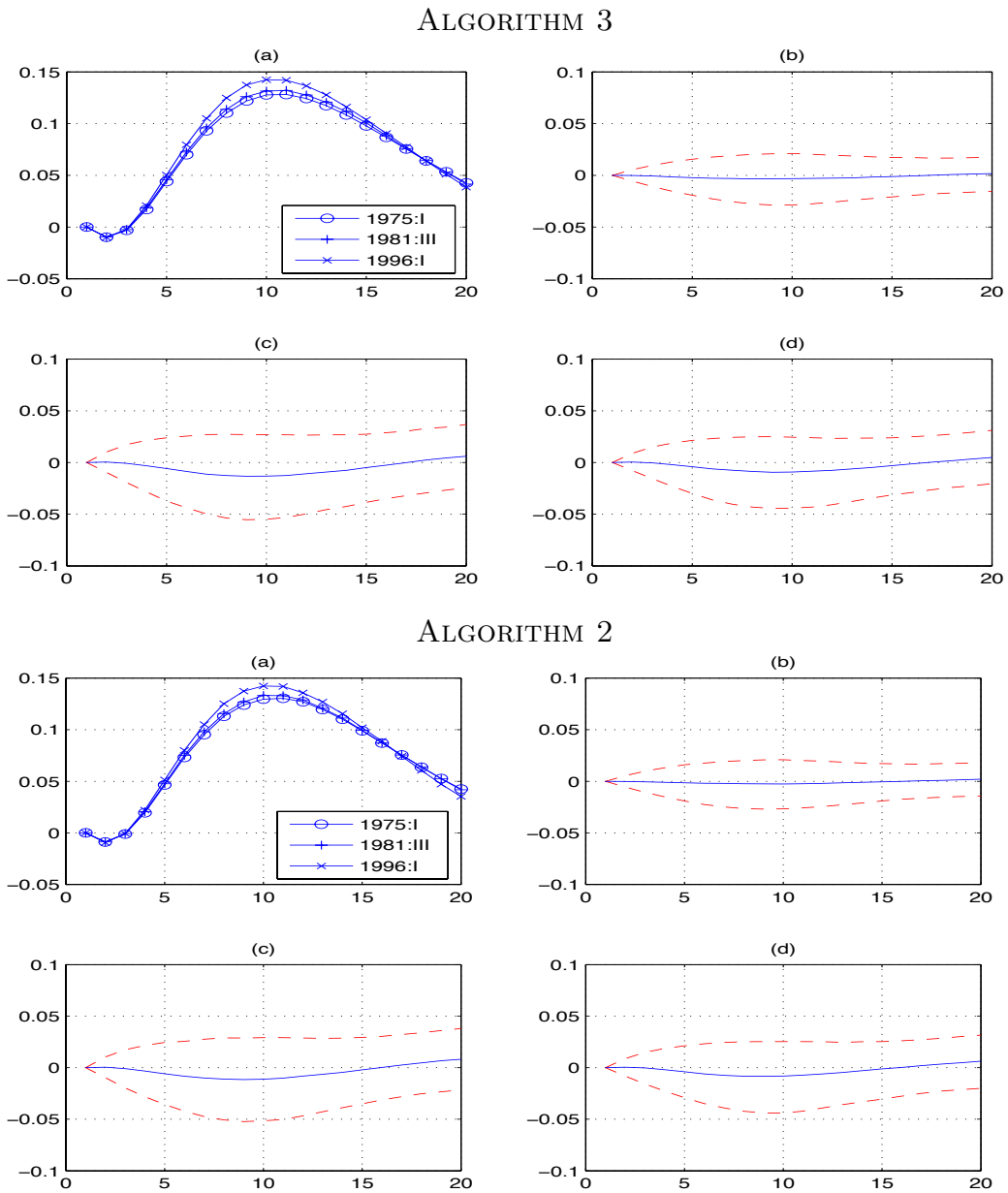

Figure 12: Interest rate response to a 1% permanent increase of inflation with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

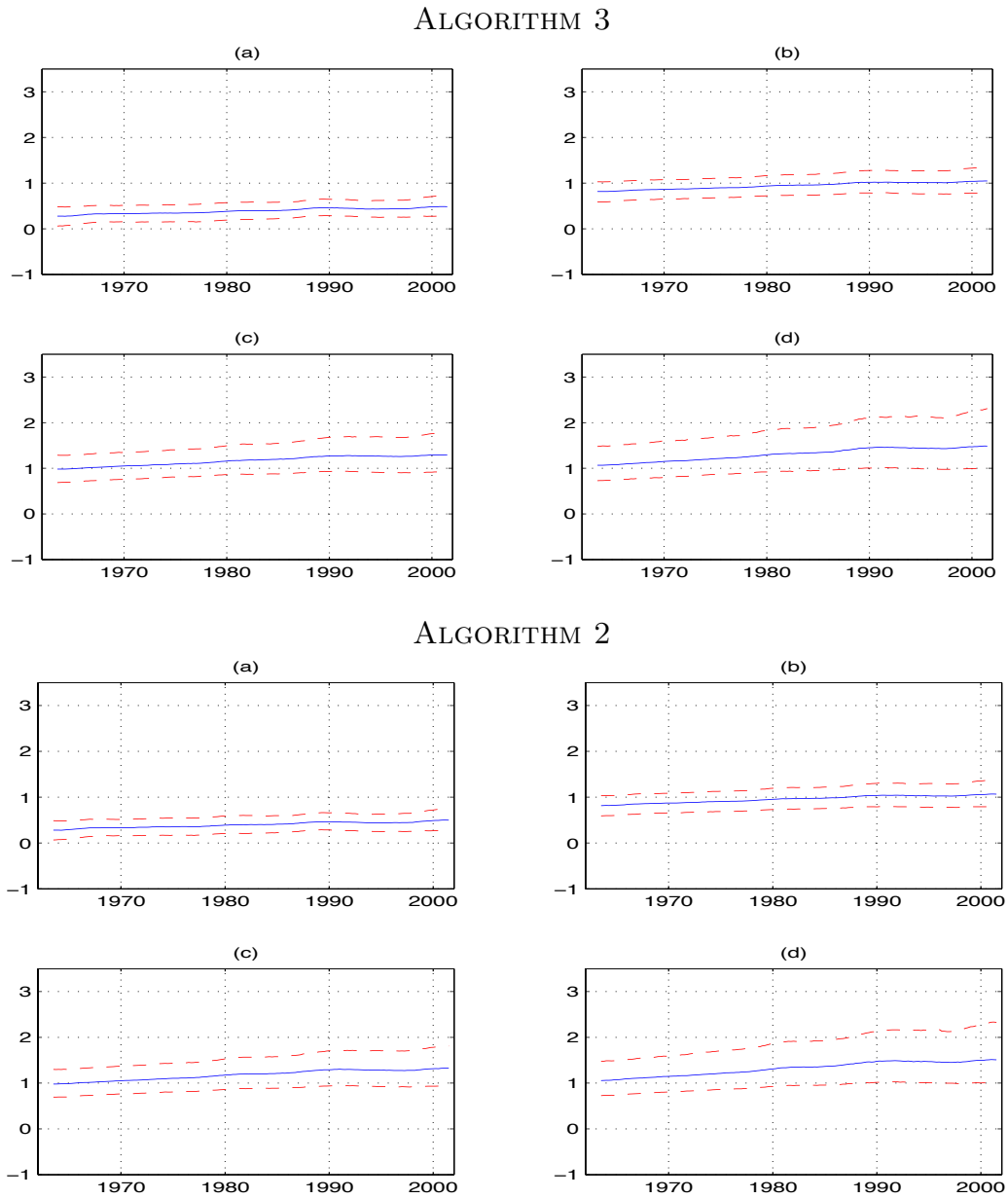

Figure 13: Interest rate response to a 1% permanent increase of unemployment with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

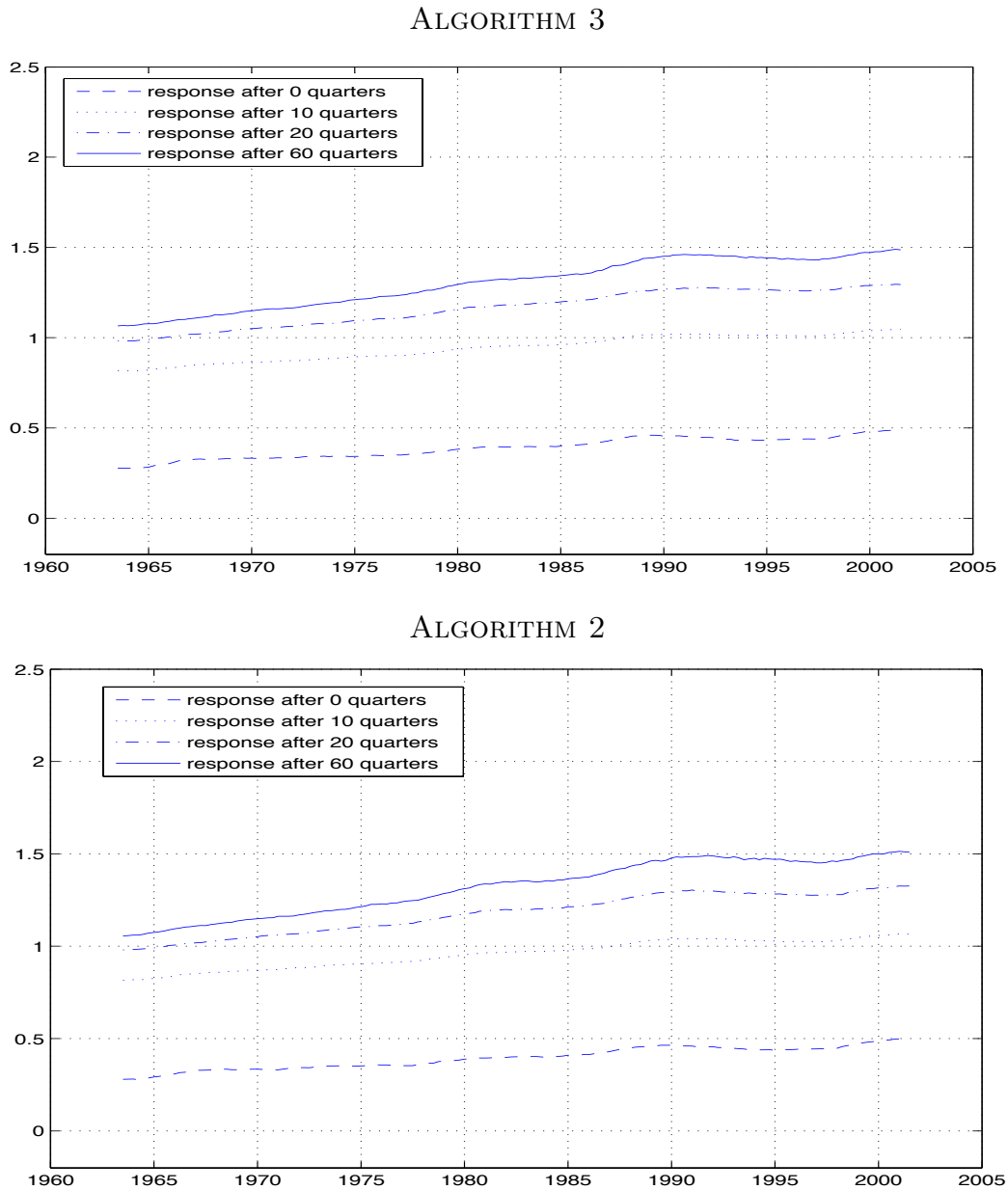

Figure 14: Interest rate response to a 1% permanent increase of unemployment with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

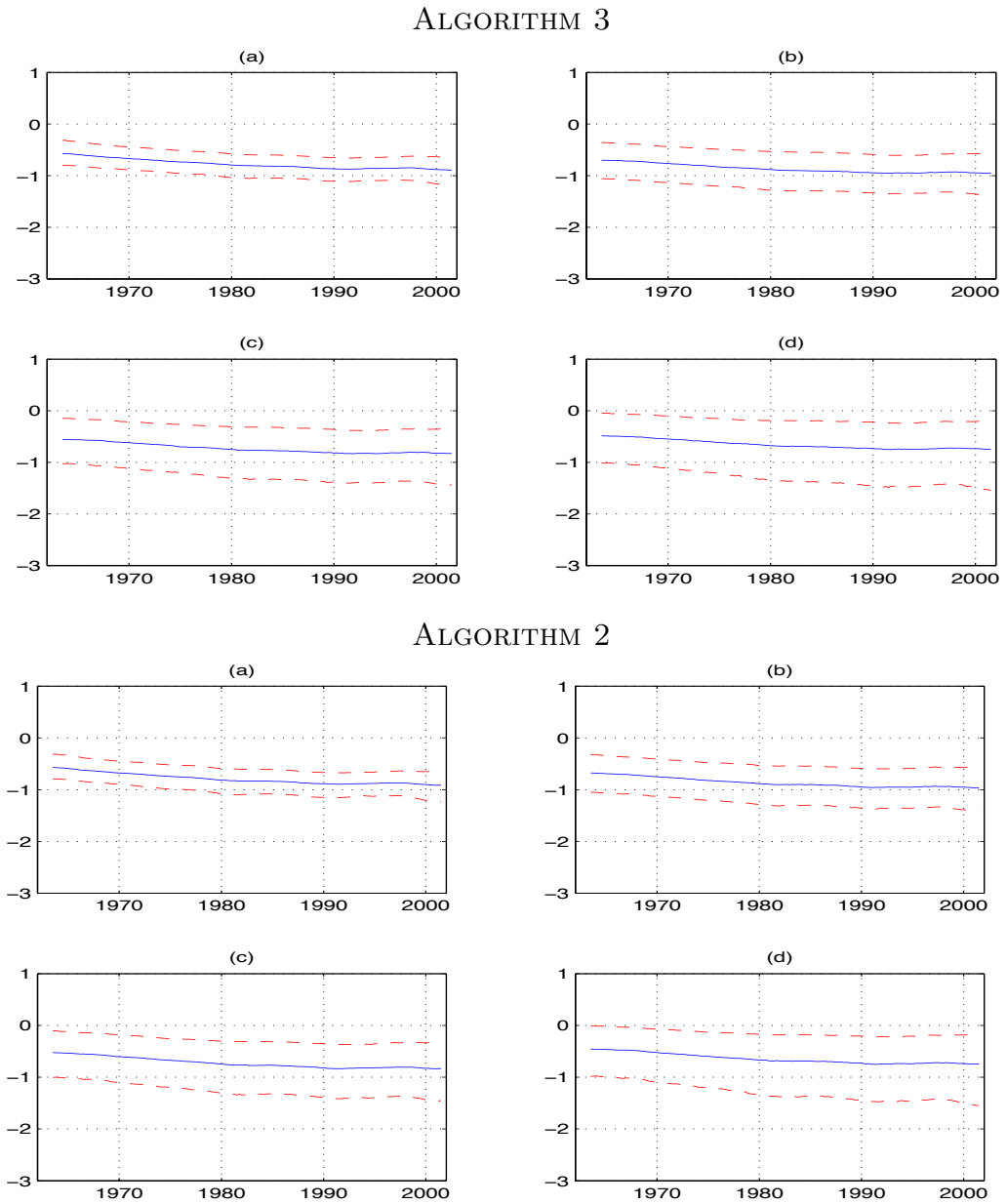

Figure 15: Interest rate response to a 1% permanent increase of unemployment.

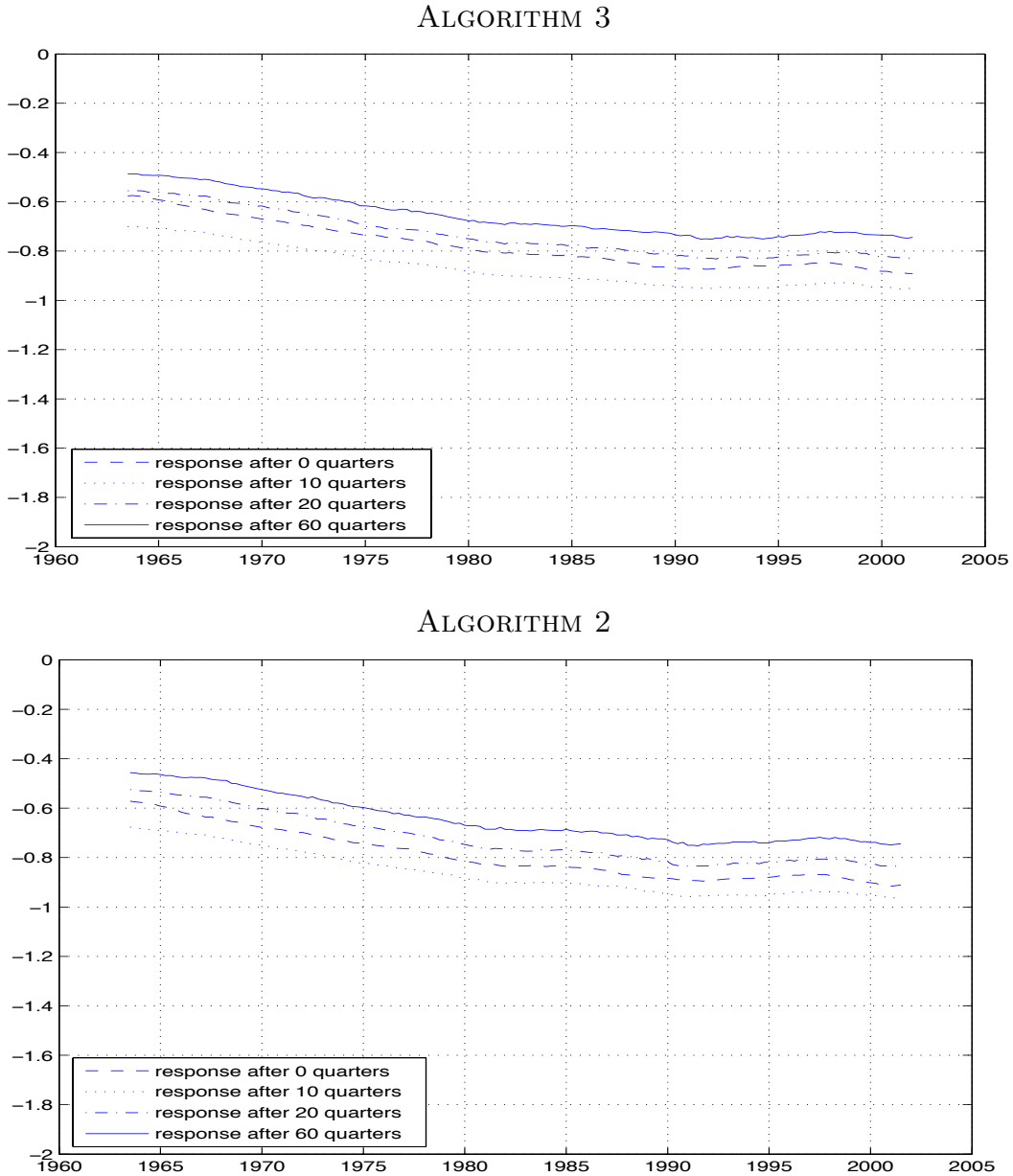

Figure 16: Counterfactual historical simulation drawing the parameters of the monetary policy rule from their 1991-1992 posterior. (a) Inflation, (b) unemployment.

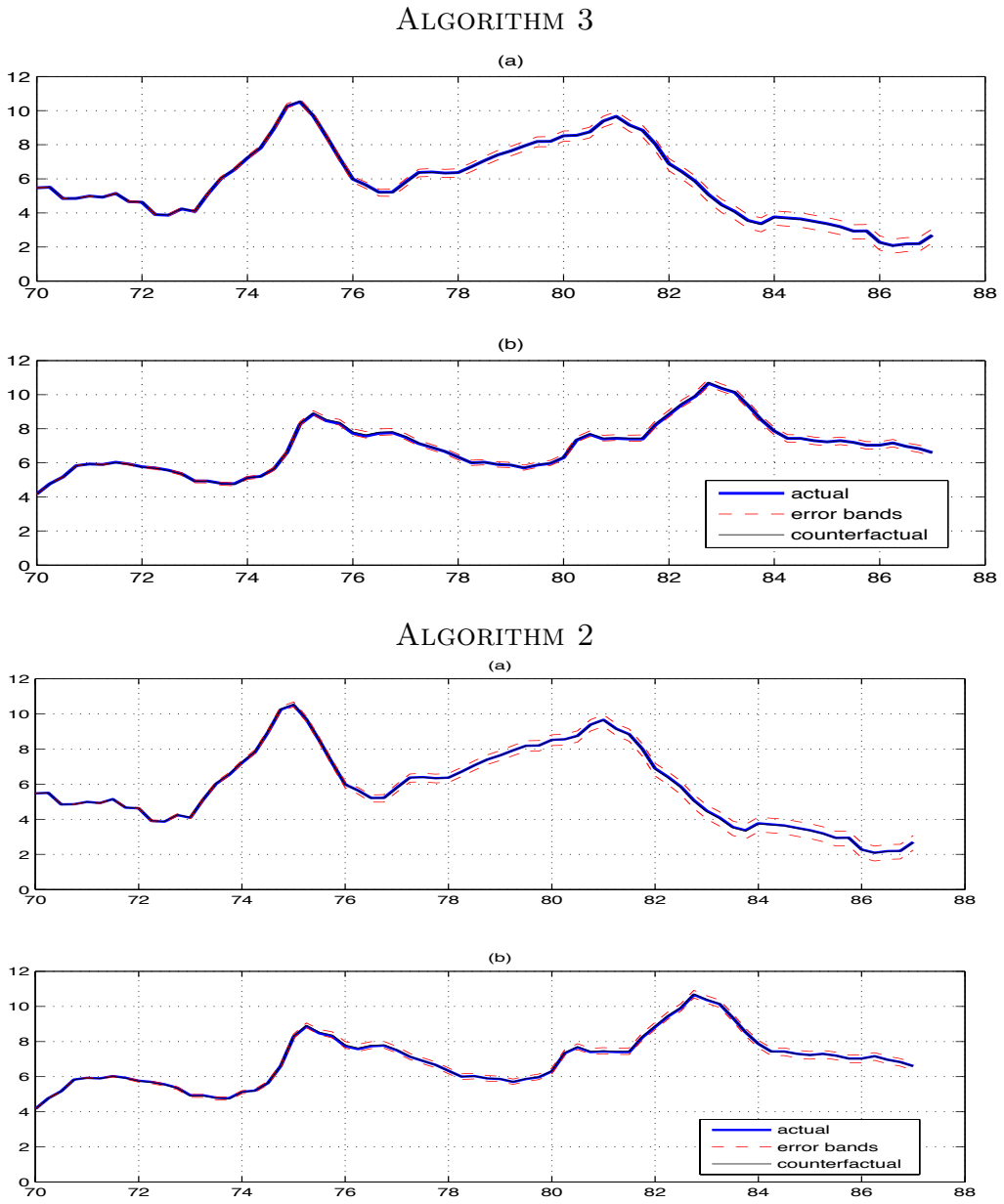

Figure 17: P-P plots obtained by applying the Geweke's (2004) procedure to Algorithm 1. The plots refer to the distribution of $\log \sigma_{i,t}$, with $t = 7$, and $i = 1$ in panel (a), $i = 2$ in panel (b), and $i = 3$ in panel (c).

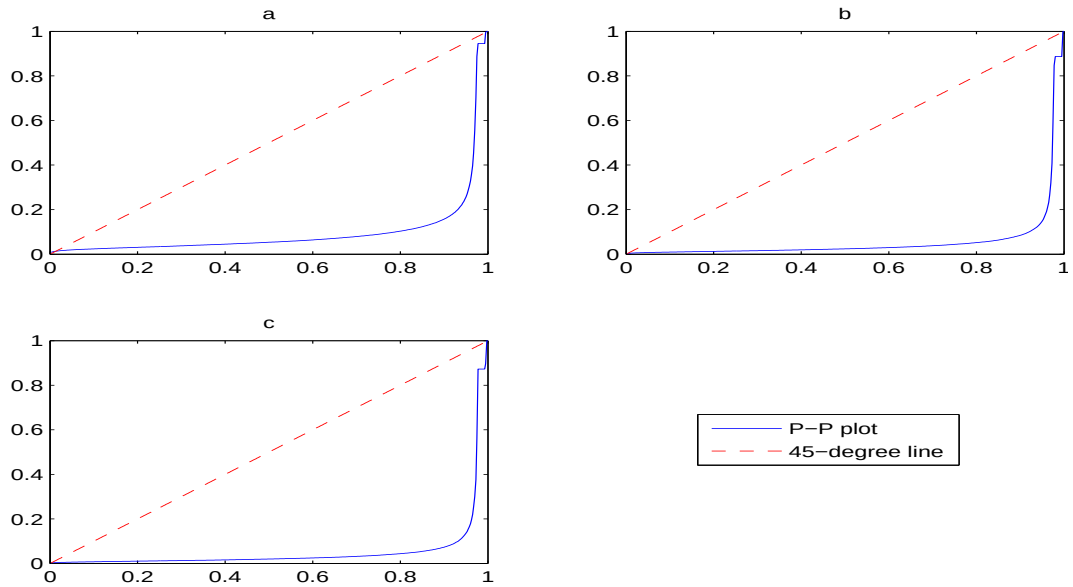

Figure 18: P-P plots obtained by applying the Geweke's (2004) procedure to Algorithm 2. The plots refer to the distribution of $\log \sigma_{i,t}$, with $t = 7$, and $i = 1$ in panel (a), $i = 2$ in panel (b), and $i = 3$ in panel (c).

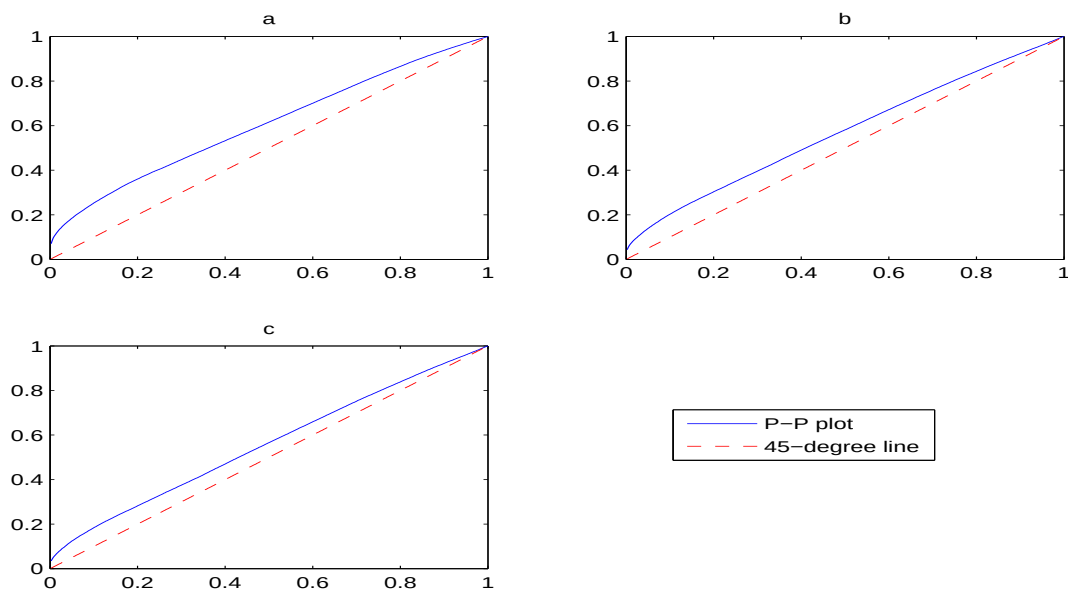

Figure 19: P-P plots obtained by applying the Geweke's (2004) procedure to Algorithm 3. The plots refer to the distribution of $\log \sigma_{i,t}$, with $t = 7$, and $i = 1$ in panel (a), $i = 2$ in panel (b), and $i = 3$ in panel (c).

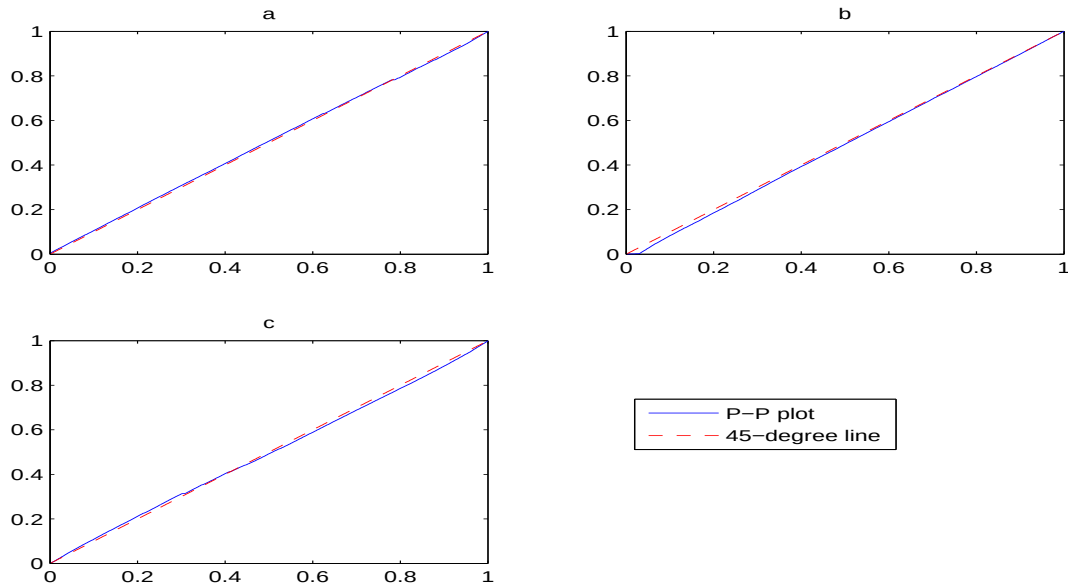

Figure 20: P-P plots obtained by applying the Geweke's (2004) procedure to Algorithm 1 augmented with a Metropolis-Hastings step to correct for the mixture-of-normals approximation error. The plots refer to the distribution of $\log \sigma_{i,t}$, with $t = 7$, and $i = 1$ in panel (a), $i = 2$ in panel (b), and $i = 3$ in panel (c).