

# Measuring Price-Level Uncertainty and Instability in the U.S., 1850-2012

Timothy Cogley\*  
Thomas J. Sargent\*\*

The views expressed herein are those of the author and do not necessarily reflect the official views of the Bank of Korea. When reporting or citing this paper, the author's name should always be stated explicitly.

---

\* New York University, [tim.cogley@nyu.edu](mailto:tim.cogley@nyu.edu)

\* New York University and Hoover Institution, [thomas.sargent@nyu.edu](mailto:thomas.sargent@nyu.edu)

For comments and suggestions, we thank Christopher Sims, Nicholas Bloom, Gabriel Mathy, Emi Nakamura, François Velde, Mark Watson, two referees, and seminar participants at Boston College and the summer 2013 SITE workshop on uncertainty and volatility. Funding from the Bank of Korea is gratefully acknowledged.

# Contents

I . Introduction .....	1
II. Data .....	4
III. A Nonlinear State-space Model with Measurement Error and Time-varying Volatilities .....	7
IV. Priors .....	9
V. Posteriors .....	11
5.1 Parameters and Latent States .....	12
5.2 Orders of Integration .....	15
5.3 Price-level Uncertainty .....	16
5.4 Price-level Instability .....	20
5.5 Deflation Risk .....	22
VI. Discussion of Related Literature .....	25
VII. Concluding Remarks .....	26
VIII. References .....	28
<Appendices> .....	30

# Measuring Price-Level Uncertainty and Instability in the U.S., 1850-2012

We use a flexible statistical model with stochastic volatilities to measure price level uncertainty and instability in the U.S. over the period 1850-2012. Major outbreaks associated with the Civil War, the two World Wars and Great Depression, and the Great Inflation and Great Recession alternate with three great price-level moderations, one near the turn of 20th century, another under Bretton Woods, and a third in the 1990s. Because periods of high and low volatility occur both before and after the Second World War, there is no evidence that the price level was systematically more stable or less uncertain in either era. Moderations sometimes involved a link to gold, but the experience of the 1990s proves that a well-managed fiat regime can achieve the same end.

**Keywords:** Price stability, inflation uncertainty, deflation risk, postwar stabilization, nonlinear signal extraction

**JEL Classification:** E31, C22

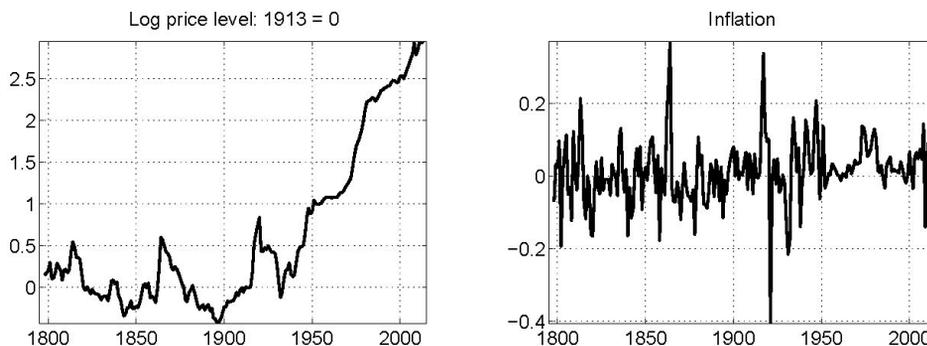
## I. Introduction

Figure 1 portrays annual data on the wholesale price level and inflation rate in the United States for the period 1798-2012. We study wholesale price indices because Christopher Hanes (2006) says they are more reliable than consumer price indices for the period before World War I. According to Hanes, statisticians collected more data on wholesale prices before 1914 than on retail prices or rents, thus compiling better raw materials for constructing price indices. Since comparability of data across periods is important for us, we examine wholesale price indices for the period after 1914 as well.

The behavior of the wholesale price level, shown in the left panel, changed dramatically after the Second World War. Between 1798 and 1940, average inflation was essentially zero, and there was no sustained growth in the price level. One very rough stylized summary of pre-World War II US monetary history is that the price level surged during wars but fell afterward, thus preserving long-run price stability. In contrast, after 1945 there was no rollback of prices, average inflation became positive, and the price level ascended. Because of that, one might suspect that price-level uncertainty was greater after World War II.

But to the extent that average inflation was predictable after World War II, a positive average would contribute to sustained growth in the price level without increasing uncertainty. The visual impression conveyed by an upward drift might therefore exaggerate the degree of uncertainty. Predictable variation would under-

Figure 1. Wholesale Price Level and Inflation



mine price level stability, however, so that the price level might have been less stable after 1945 even if were no more uncertain.

In addition, although average inflation was zero before 1940, there was a lot of volatility. Notice in particular that the rollbacks after the War of 1812 and the Civil War lasted for decades and eventually generated more cumulated postwar deflation than wartime inflation.<sup>1)</sup> Uncertainty about the extent and duration of a rollback especially about the degree of overshooting – would at least partially offset its presumed benefits. Furthermore, data on inflation, exhibited in the right panel, suggest that volatility might have been lower *after* 1950, despite the Great Inflation and 2008 financial crisis. For these reasons, whether price-level uncertainty increased or decreased after 1945 is unclear.

In this paper, we estimate a flexible statistical model for wholesale price inflation and use it measure price-level uncertainty and instability before and after World War II. Our starting point is a simple and elegant model for inflation developed by Stock and Watson (2007). A version of their model forms the state equations for a nonlinear state-space representation. Motivated by the work of Christina Romer (1986a,b), we also append measurement error to cope with noise in pre-World War II data. As Romer points out, one of the chief obstacles one must face when assessing volatility across historical eras is that data are not directly comparable. Long historical time series such as the one shown in figure 1 are typically spliced together from a variety of sources that define and measure objects differently. Older data are usually noisier, and measurement error amplifies volatility.

To level the playing field, Romer posed and solved a ‘noise-creation’ problem, constructing a comparably noisy series for the postwar era by imitating prewar measurement procedures using postwar data sources. After estimating the variances of comparably noisy measures, she concluded that much of the apparent decline in U.S.

unemployment and GDP volatility is a consequence of better measurement. Similarly, Hanes (1998) solved a noise-creation problem for U.S. wholesale price

---

1) For instance, during the deflation following the Civil War, the price level overshot its 1860 value on the way down and eventually fell 30 percent below the prewar mark. After reaching a trough in 1896, another 10 years passed before the price level returned to its 1860 value.

indices, constructing a noisy post-World War II series whose stochastic properties are consistent with prewar data.

This paper extends Hanes's analysis by solving a signal extraction problem for prewar data, thus improving estimates of historical volatility. We begin by describing a nonlinear state-space model in which prewar inflation is assumed to be measured with error. So that Romer's conclusion is not hardwired, our model allows true but hidden volatilities to vary over time, as in Stock and Watson (2007). We solve the signal extraction problem by applying Bayes's theorem. If the state-space model were linear and conditionally Gaussian, this would reduce to an operation closely related to Kalman smoothing. Because the model is nonlinear in state variables, however, we use a Markov Chain Monte Carlo (MCMC) algorithm to simulate the posterior distribution for the model's parameters and hidden variables. From that posterior, we deduce the posterior distribution for conditional variances and second moments of future price levels, and we compare our measures of uncertainty and instability at various dates.

Specification of the measurement-error process and identification of its parameters are key elements of our analysis. Because measurement error is a latent variable weakly restricted by theory, how to specify it and how to identify its parameters are not obvious. Here we exploit Hanes's consistent postwar measure of wholesale price inflation. For the period 1948-1990, Hanes's noisy measure overlaps with the putatively noise<sup>2)</sup> free measure constructed by the Bureau of Labor Statistics. The availability of two inflation measures – one noisy and one clean – sharply identifies measurement error for years in which they overlap, allowing precise estimation of its parameters.

After solving this signal extraction problem, we find that uncertainty was highest and the price level least stable during the Civil War, the two World Wars and interwar years, and at the time of the Great Inflation and Great Recession. There were also three great price-level moderations, one near the turn of the 20th century, another under Bretton Woods, and a third during the 1990s. Because periods of high and low volatility occur both before and after the Second World War, we

---

2) At least this is the conventional assumption in macroeconomics.

find no compelling evidence that the price level was systematically more stable or less uncertain in either era. There was less uncertainty and more stability during moderations than at other times, but moderations are found in both periods.

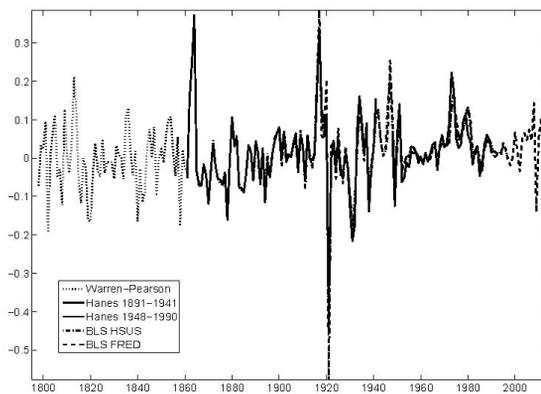
## II. Data

Our data are annual and are assembled from four sources, Warren and Pearson (1933), Hanes (1998), and two series from the Bureau of Labor Statistics (*BLS*). Hanes's (2006) chapter in the Millennial edition of the *Historical Statistics of the United States* (*HSUS*) describes the historical data and documents its sources, and the data itself can be found in the accompanying tables.<sup>3)</sup> Postwar data were downloaded from the *Federal Reserve Economic Database* (*FRED*)

Data for the period 1797-1890 come from Warren and Pearson (1933). The dotted line in figure 2 depicts their measure of inflation. We use this series primarily as a training sample to calibrate aspects of the prior.

Hanes (1998) replicated and extended Warren and Pearson's series for the periods 1860-1941 and 1947-1990 (see the solid line in figure 2). His objective was not to construct the best measure of wholesale price inflation for these periods but to

Figure 2. US Wholesale Price Inflation:  
Overlapping Sources



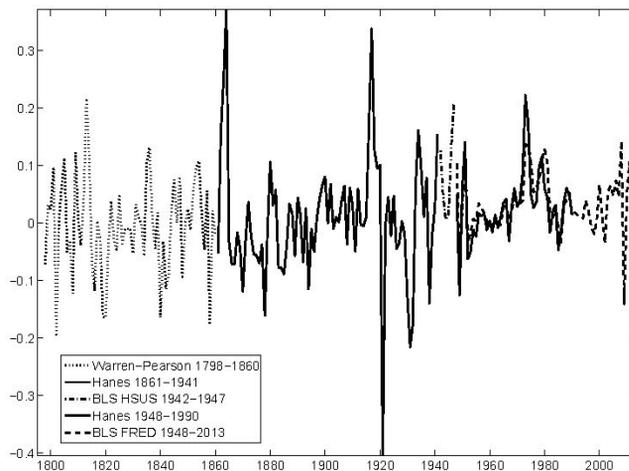
3) See <http://hsus.cambridge.org/HSUSWeb/toc/showChapter.do?id=Cc>.

follow Warren and Pearson as closely as possible, thus constructing a series consistent with theirs. For years in which their measures overlap, the correlation is 0.997, verifying that Hanes's procedures are faithful to those of Warren and Pearson.<sup>4)</sup> Among other things, we interpret 'consistency' to mean that the stochastic properties of measurement errors in Hanes's data are the same before and after the Second World War.

The United States government began constructing wholesale price indices in 1902 and produced retrospective measures back to 1890. We make use of two data series from the *BLS*, one from *HSUS* covering the period 1890-1997 and another from *FRED* covering the period 1913-2013. The corresponding inflation measures are shown as dashed and dashed-dotted lines, respectively, in figure 2. The two *BLS* series are very highly correlated for the period in which they overlap (correlation 1914-1997 = 0.986), but because *FRED* data were revised more recently we believe they are more accurate for the period after the Second World War. Differences are slight in any case.

Although the Warren-Pearson-Hanes series is our primary noisy measure, there is a gap from 1942-1947 that we fill with *BLS-HSUS* data. That the *BLS-HSUS*

Figure 3. US Wholesale Price Inflation: Spliced Sources



4) Indeed, for the overlapping years, Warren and Pearson's dotted line in figure 2 hides behind Hanes's solid line.

measure is comparable to the Warren-Pearson-Hanes series is not obvious, but the following defense can be offered. Warren and Pearson's objective was to construct a retrospective measure comparable to the *BLS* index begun in 1890, and Hanes's goal was to mimic Warren and Pearson (see Hanes 2006). To the extent that both succeeded, one might therefore expect early *BLS-HSUS* data to correlate highly with Hanes's measure. Indeed, for the period 1914-1941, the coefficient in a projection of Hanes's inflation measure onto that of *BLS-HSUS* is 1.02 with a standard error of 0.04, and the correlation is 0.978. Thus, although some discrepancies emerge, they seem minor. Extrapolating this close association to 1942-1947 rationalizes bridging the gap between Hanes's two series with *BLS-HSUS* data.<sup>5)</sup>

After discarding redundant data, we are left with the data shown in figure 3 that we will use to calibrate priors and estimate the model. We construct a noisy inflation measure for 1798-1990 by splicing data from Warren-Pearson-Hanes and *BLS-HSUS*.<sup>6)</sup> Our clean post-World War II measure (1948-2012) comes from *BLS-FRED*. Thus two inflation measures – one noisy and one clean – are available for 1948-1990, and a single measure – either noisy or clean – is available for all other years. That a noise ridden series comparable to that for the period before World War II overlaps with a clean measure for much of the postwar period underlies our strategy for coping with measurement error.

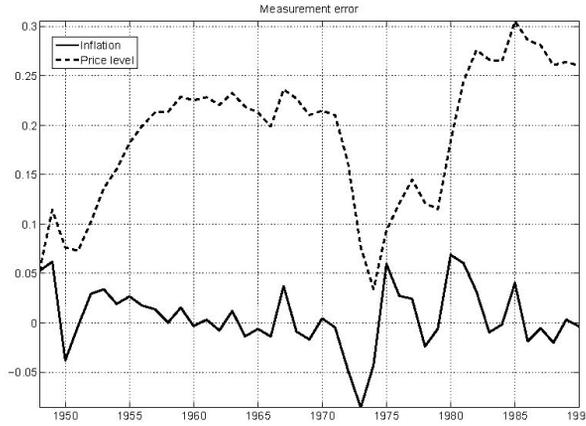
For guidance on how to represent measurement error, we examine the difference between the noisy and clean measures  $y_{Hanes} - y_{BLS-FRED}$  for the period 1948-1990. On the assumption that *BLS-FRED* data are error free, this difference identifies noise in Hanes's measure of inflation, and its partial sum identifies noise in his measure of the log price level. As shown in figure 4, noise in inflation (depicted by the solid line) appears to be stationary, centered near zero, and weakly autocorrelated, while noise in the log price level (shown as a dashed line) appears to have a unit root. Thus, in what follows, we assume that our noisy measure of inflation is the sum of true inflation plus a mean zero, stationary measurement error.

---

5) We also estimated a version of our model that treats this gap as missing, and the conclusions were essentially the same.

6) In principle, precision could be enhanced by using all the noisy data, but since the various noisy measures are highly correlated we suspect the gain would be slight. Thus we simplify by using a single noisy series.

Figure 4. Measurement Error in Inflation and the Log Price Level, 1948-1990



### III. A Nonlinear State-space Model with Measurement Error and Time-varying Volatilities

Following Stock and Watson (2007) and Shephard (2013), the state equations for our statistical model are

$$\begin{aligned}
 \pi_t &= \mu_t + \sqrt{r_t} \epsilon_{\pi t} \\
 \mu_t &= \mu_{t-1} + \sqrt{q_t} \epsilon_{\mu t} \\
 \ln r_t &= \ln r_{t-1} + \eta_{rt} \\
 \ln q_t &= \ln q_{t-1} + \eta_{qt}
 \end{aligned} \tag{1}$$

where  $\pi_t$  is inflation,  $\mu_t$  is trend inflation, and  $r_t$  and  $q_t$  are stochastic volatilities that evolve as geometric random walks. The innovations  $\epsilon_{\pi t}$ ,  $\epsilon_{\mu t}$  are standard normal, serially uncorrelated, independent of all the other shocks in the model. Similarly, the log volatility innovations  $\eta_{rt}$  and  $\eta_{qt}$  are iid normal with mean zero and covariance matrix

$$W = \begin{bmatrix} \sigma_r^2 & \sigma_{rq} \\ \sigma_{rq} & \sigma_q^2 \end{bmatrix} \tag{2}$$

This parameterization was introduced by Shephard (2013) and extends that of Stock and Watson (2007) by allowing correlated increments to  $\ln q_t$  and  $\ln r_t$ . Shephard reports that this extension improves the model's fit to quarterly post World War II US inflation.<sup>7)</sup>

This model has three features that we believe are necessary for an adequate statistical representation of inflation. First, to fit both the 19th century, when average inflation was close to zero, and the period after World War II, when average inflation was positive and variable, the model must include a stochastic trend. There must also be a transient component to fit the short-term volatility seen at various times throughout the sample. Third, so that volatility can change, innovations to the two components must have time-varying variances.

Our main extension of the Stock-Watson-Shephard model is to confront the measurement issues raised by Romer. To address her concern, we append an additional state equation for a serially correlated measurement error  $m_t$ ,

$$m_t = \rho_m m_{t-1} + \sigma_m \epsilon_{mt} \quad (3)$$

where  $\epsilon_{mt}$  is iid standard normal and independent of the other shocks in the model. The measurement equation for our state-space model is

$$y_t = C_{mt} \begin{bmatrix} \pi_t \\ m_t \end{bmatrix} \quad (4)$$

where  $y_t$  represents observations on inflation. Our identifying assumptions are that the post-1948 *BLS-FRED* data are noise free and that the measurement error parameters  $\rho_m$  and  $\sigma_m$  are constant throughout Hanes's sample. This is how we interpret the statement that Hanes's pre- and postwar measurement procedures are consistent. Under these assumptions, the measurement arrays  $y_t$  and  $C_{mt}$  are defined as

◦  $y_t = y_{nt}$ ,  $C_{mt} = [1 \ 1]$  when a single noisy observation  $y_{nt}$  is available (before 1948);

7) As we shall see, we find less evidence of correlation in log-volatility innovations, perhaps because our data are annual, measure wholesale prices instead of consumer prices, cover a longer time span, and are noisier than Shephard's. We also estimated a version based on Stock and Watson's original parameterization that constrains  $\sigma_{r,d} = 0$  and obtained results that are similar to those reported below.

- $y_t = y_{ct}$ ,  $C_{mt} = [1 \ 0]$  when a single clean observation  $y_{ct}$  is available (after 1990);
- $y_t = [y_{nt} \ y_{cy}]'$ ,  $C_{mt} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$  when both noisy and clean observations are available (1948-1990).

## IV. Priors

Our first task is to estimate the latent states  $\pi_t$ ,  $\mu_t$ ,  $r_t$ ,  $q_t$ , and  $m_t$ , the log-volatility innovation variance  $W$ , and the measurement error parameters  $\rho_m$  and  $\sigma_m$ . We do this via Bayesian methods.

Priors are wanted for the initial states  $\mu_0$ ,  $\pi_0$ ,  $q_0$ ,  $r_0$ ,  $s_0$ , and  $m_0$  as well as the parameters  $W$ ,  $\rho_m$ , and  $\sigma_m$ . The transition equations then imply priors for the remaining states  $\mu_t$ ,  $\pi_t$ ,  $q_t$ ,  $r_t$  and  $m_t$ . We assume priors for  $\mu_0$ ,  $\pi_0$ ,  $r_0$ ,  $q_0$ ,  $m_0$ ,  $W$ ,  $\rho_m$ , and  $\sigma_m$  to be mutually independent, and specify marginal priors for each. Many aspects of the prior are calibrated using a training sample consisting of measured inflation for 1798-1849.

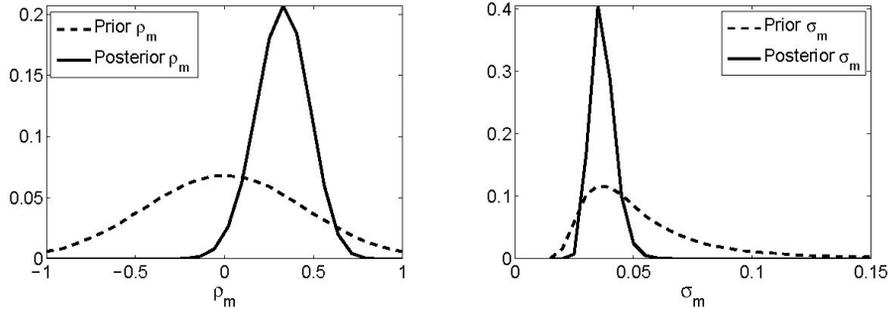
For instance, the prior for the initial state  $(\mu_0, \pi_0)$  is normal with mean equal to the training sample average (-0.90 percent per annum) and variance

$$P_0 = \begin{bmatrix} 0.15^2 & 0 \\ 0 & 0.025^2 \end{bmatrix} \quad (5)$$

Since prior credible sets for  $\pi_0$  and  $\mu_0$  are roughly (-0.3,0.3) and (-0.05,0.05), respectively, the prior is weakly informative about initial inflation.

We also adopt normal priors for  $\ln r_0$  and  $\ln q_0$ , the logs of the initial innovation variances for the transitory and persistent components of inflation, respectively. The prior mean for  $\ln r_0$  is the log of the training sample variance for measured inflation (-4.97), while that for  $\ln q_0$  is the log of the training sample variance divided by 25 (-8.19). We set the prior standard deviation for both to 5, a value that is huge on a log scale. This makes the prior on the initial values  $\ln r_0$  and  $\ln s_0$  very weakly informative.

Figure 5. Priors and Posteriors for Measurement Error Parameters



For the measurement error parameters  $\rho_m$  and  $\sigma_m$ , we adopt weakly informative priors that are portrayed by dashed lines in figure 5. The prior for  $\rho_m$  is normal with mean zero and standard deviation 0.45, thus centering on a white-noise specification and concentrating the preponderance of prior mass in the stationary region.<sup>8)</sup> For  $\sigma_m$ , we adopt an inverse-gamma prior whose mode equals 50 percent of the training sample standard deviation for measured inflation (70.71 percent of the variance), thereby encoding an initial belief that the Warren-Pearson-Hanes data are very noisy. Not wanting to hardwire this belief, however, we set the prior degrees of freedom to 2, so that a centered 95 credible set ranges from 28 percent to more than 100 percent of the training sample standard deviation. The result is an  $IG_1(0.051, 2)$  specification.

A prior for  $W$ , the covariance matrix for log-volatility innovations, is harder to calibrate using the training sample. Instead, we adopt an informative inverse-Wishart prior that is inspired by Stock and Watson's calibration. We start with their parameter for the variance of log-volatility innovations in quarterly data, adjust for time aggregation to an annual sampling frequency, and then set the diagonal elements of the prior scale matrix so that the prior modes for  $\sigma_r^2$  and  $\sigma_q^2$  equal the adjusted value. Lacking a strong prior view about the covariance  $\sigma_{rq}$ , we set the off-diagonal elements of the prior scale matrix to zero, thus centering the prior covariance on zero. After centering the prior in this way, we set the degree of free-

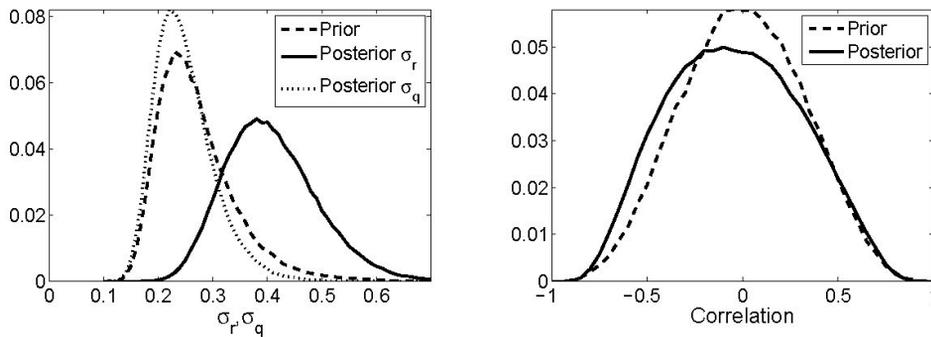
8) Truncating the prior at the boundary of the stationary region was unnecessary because all the posterior draws were in the interior.

dom parameter to deliver plausible prior credible sets. After some experimentation, we settled on an  $W(0.05 \cdot I_2, 10)$  prior for  $W$ . Dashed lines in figure 6 portray prior histograms for the standard deviations  $\sigma_r$  and  $\sigma_q$  and correlation  $corr_{rq}$ . A centered prior 95 percent credible set for the standard deviations ranges from 0.158 to 0.380, while that for the correlation covers the interval  $\pm 0.6$

## V. Posteriors

The joint posterior distribution for parameters and hidden states is simulated via a MCMC algorithm. Except for the measurement error and assumptions about the prior, the model is very similar to that of Stock and Watson (2007) and is closely related to that of Cogley and Sargent (2005). The MCMC algorithm is therefore also similar to Cogley and Sargent's. Appendix A describes the details.<sup>9)</sup>

Figure 6. Priors and Posteriors for Variances and Covariances of Log-Volatility Innovations



<sup>9)</sup> Shephard (2013) describes an alternative simulation algorithm based on an auxiliary particle filter. Because his algorithm intensively exploits parallel processing, our guess is that it is computationally more efficient than ours. Nevertheless, our algorithm delivers a valid sample from the joint posterior distribution for parameters and hidden states.

## 5.1 Parameters and Latent States

As we shall see, our measures of uncertainty and instability are functions of model parameters and latent states. We therefore begin by summarizing aspects of their posterior distribution.

Also shown in figure 5 are posterior distributions for the measurement-error parameters  $\rho_m$  and  $\sigma_m$ . As expected, measurement-error parameters are sharply identified from overlapping clean and noisy observations on inflation, and their posteriors are much more concentrated than the priors (compare the solid and dashed lines in figures 5). Measurement errors are autocorrelated (the mode for  $\rho_m$  is 0.33, and a centered 95 percent credible set ranges from 0.03 to 0.6), and there is little support for large values of  $\sigma_m$  (the posterior mode is 0.035, and a centered 95 percent credible set ranges from 0.028 to 0.045). At the posterior mode, the unconditional standard deviation of  $m_t$  is 0.037. Since the standard deviation of the Warren-Pearson-Hanes measure of inflation is 0.092, this amounts to 40 percent of measured inflation volatility, a substantial proportion but considerably less than some values entertained by the prior.

Posterior distributions for the standard deviations and correlation of log volatility innovation appear in figure 6. Perhaps less obviously, precise estimation of measurement error parameters helps identify pre-1948 movements in  $r_t$  and therefore also improves estimation of  $\sigma_r$ . The noise-ridden measure of inflation has two transient components, one that is genuine  $\sqrt{r_t} \varepsilon_{\pi t}$  and another due to measurement error  $m_t$ . Since parameters governing the latter are precisely estimated, the amount of transient variation attributable to measurement error is restricted. Transient variation over and above an amount consistent with  $m_t$  must therefore be genuine. Since pre-1948 variation in  $r_t$  is more strongly identified, so is  $\sigma_r$ . Thus, in contrast to our earlier work (Cogley et al. 2012), the posterior for  $\sigma_r$  differs from its prior, with the data wanting higher values of  $\sigma_r$  (compare the solid and dashed lines figure 6). The posterior mode for  $\sigma_r$  increases to 0.38, and a centered 95 percent credible set ranges from 0.26 to 0.59.

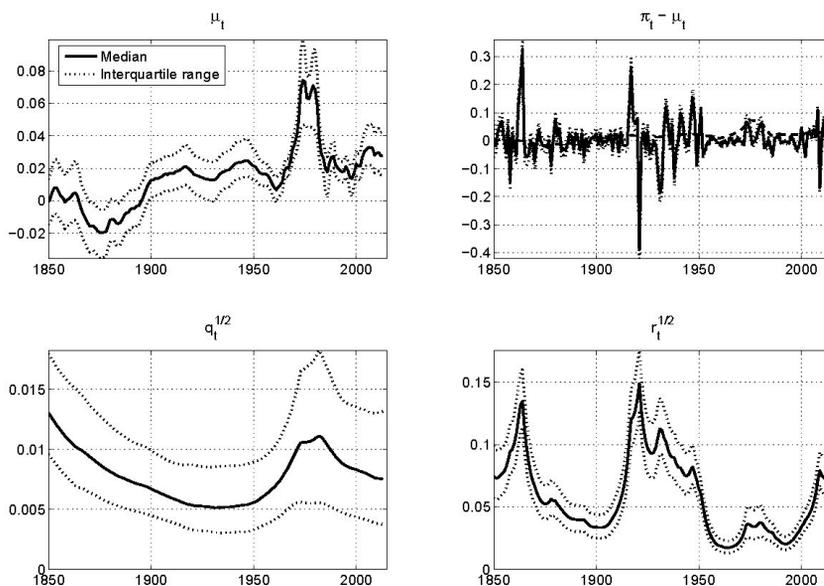
Alas, the other innovation variance parameters are weakly identified, and their

posteriors are essentially the same as their priors (compare the dotted and dashed lines in the left panel of figure 6 and the solid and dashed lines in the right). Because of the extent to which the prior and posterior for  $\sigma_q$  overlap, we also examined a model with a less informative  $W(0.05 \cdot I_2, 2)$  prior for  $W$ . This has long upper tails for  $\sigma_q$  and  $\sigma_r$  and thus promotes faster drift in  $q_t$  and  $r_t$ . Despite that, the posteriors were similar to those shown in figure 6, and our overall conclusions were much the same. For this reason, we doubt that our benchmark prior has undue influence on the results.

Aspects of the posterior for the hidden states  $\mu_t$ ,  $\pi_t - \mu_t$ ,  $\sqrt{q_t}$ , and  $\sqrt{r_t}$  are shown in figure 7. In each panel, a solid line depicts the posterior median at each date, and dotted lines portray the interquartile range.

The persistent component  $\mu_t$  was centered around zero in the 19th century but was not strongly anchored near zero (see the upper left panel). On the contrary, the posterior median wandered between +80 and -200 basis points per annum until the early 20th century, when it settled in the vicinity of 1.5 to 2 percent. It remained there until the Great Inflation of the 1970s, when the median estimate peaked at 7.4 percent per annum. After the Volcker disinflation, the median estimate fell to

Figure 7. Posteriors for Hidden States



around 2.5 percent, and it fluctuated between 1.4 and 3.3 percent from then until the end of the sample. There is also substantial uncertainty about the location of  $\mu_t$  at any given date. Before 1914, the interquartile range was typically  $\pm 130$  basis points around the median with a spike to  $\pm 180$  basis points during the Civil War. With better data after 1948, the interquartile range narrowed to  $\pm 85$  basis points around the median but spiked to  $\pm 260$  basis points during the Great Inflation.

Movements in  $q_t$ , the innovation variance of the persistent component  $\mu_t$ , follow a similar pattern (see the lower left panel of figure 7). Volatility was high during the third quarter of the 19th century and declined gradually, with the median estimate of  $\sqrt{q_t}$  falling from 1 percent in 1865 to 0.5 percent in 1945. As the economy's nominal anchor began to drag during the run-up to the Great Inflation, our estimate of  $\sqrt{q_t}$  rose, reaching a local peak of 1.1 percent at the time of the Volcker disinflation. Thus, the post-World War II peak in  $q_t$  was about the same as at the end of the Civil War. Furthermore, the median estimate fell after 1983 and by the end of the sample had returned to roughly the same level as at the end of the 19th century. We therefore find no strong evidence that increments to the persistent component of inflation were systematically more variable after World War II. On the contrary, the volatility of  $\Delta \mu_t$  was highest at the beginning of our sample.

The right column of figure 7 depicts the transient component of inflation  $\pi_t - \mu_t$  and its innovation variance  $r_t$ . Although perhaps not obvious at first glance, the figure attests that most bursts in U.S. inflation were transitory. To see this, notice the difference in scale between the left and right columns. For instance, while the interquartile range for  $\mu_t$  varies between -0.03 and 0.09, that for  $\pi_t - \mu_t$  ranges from -0.4 to 0.4. Similarly, estimates of  $\sqrt{r_t}$  are an order of magnitude greater than those for  $\sqrt{q_t}$ . To highlight this difference in scale, we reproduce in the top right panel the median estimate of  $\mu_t$ , which we portray as a dashed line. Except for the Great Inflation, fluctuations in  $\mu_t$  are small in comparison with variations in  $\pi_t - \mu_t$ . As we shall see, transient variation dominates not only at the one-year horizon but also at longer horizons.

The largest bursts of transient inflation occurred during and after the Civil War, in and between the two World Wars, and during the recent financial crisis. A lesser

burst of transient volatility also appeared during the Great Inflation, but this was comparatively small. Troughs in  $r_t$  occurred shortly after the turn of the 20th century, in the early 1960s under Bretton Woods, and in the 1990s. Thus, three ‘great moderations’ appear in our sample, one in the last quarter of the 19th century, another between World War II and the Great Inflation, and a third beginning in the mid 1980s and ending after the turn of the 21st century. Because measurement error has been purged, the model asserts that these moderations are genuine.

## 5.2 Orders of Integration

Our model implies that inflation is  $I(1)$  and that the log price level is  $I(2)$ , a common specification for post-World War II data. However, whether these orders of integration are consistent with data from before the First World War data is not obvious. To examine this issue, we estimate an augmented Dickey-Fuller regression,

$$y_t = \mu + \rho y_{t-1} + \sum_{j=1}^2 \zeta_j \Delta y_{t-j} + u_t \quad (6)$$

where  $y_t$  is measured inflation for the period 1850-1913, and we calculate the  $t$ -statistic for  $\rho - 1$ . This augmented Dickey-Fuller statistic is -4.56, and its 1-percent asymptotic critical value is -3.43. The test therefore seems strongly to reject a unit root in inflation.

However, the estimates shown in figure 7 suggest that the random-walk component of inflation was small prior to 1914 and that the transient component plus measurement error was much more volatile. It is tenuous whether an augmented Dickey-Fuller test can detect a small random-walk component hidden under substantial noise. To check the size of the test for our specification, we simulate the state-space model, generating artificial data on measured inflation by drawing from the posterior for  $\pi_t$ ,  $\rho_m$ , and  $\sigma_m$  and then calculating the implied distribution for the augmented Dickey-Fuller statistic. It turns out that the null distribution is shifted well to the left of the asymptotic distribution and that the correct 10 percent critical value is -5.13, implying that a unit root in inflation is not rejected. Indeed, the  $p$ -value for a sample statistic of -4.56 is 0.39.

None of this proves that inflation was  $I(1)$ , but it does establish that our  $I(1)$  representation is not grossly at odds with the data. We are sufficiently reassured

that our representation for inflation is good enough for pre-World War I data to allow us to proceed.

### 5.3 Price-level Uncertainty

To measure price-level uncertainty, we calculate the standard deviation of cumulative  $h$ -year inflation,  $p_{t+h} - p_t$ , conditioned on the log-volatility innovation variance  $W$  and histories of the time-varying states  $w^t = (\pi^t, \mu^t, r^t, q^t)$ . As shown in appendix B, the conditional mean and standard deviation are

$$\begin{aligned} E(p_{t+h} - p_t | w^t, W) &= h\mu_t \\ \sigma(p_{t+h} - p_t | w^t, W) &= \sqrt{q_t \sum_{j=1}^h (h-j+1)^2 \exp(j\sigma_q^2/2) + r_t \sum_{j=1}^h \exp(j\sigma_r^2/2)} \end{aligned} \quad (7)$$

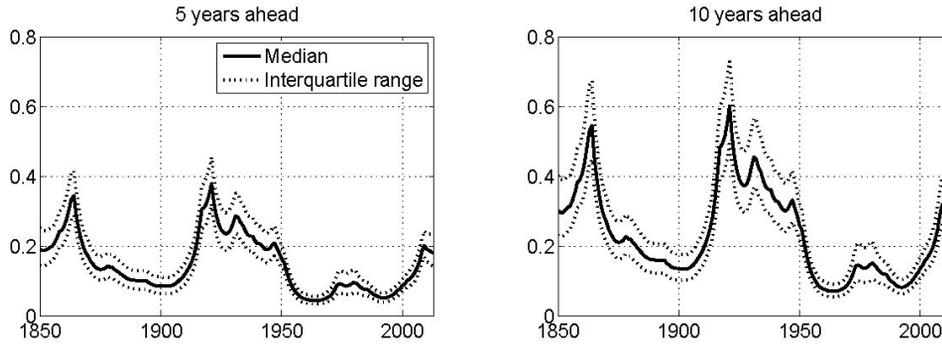
respectively. Because the information set contains the current price level  $p_t$ , the conditional standard deviation of the future price level  $p_{t+h}$  is the same as that for cumulative inflation,  $\sigma(p_{t+h} | w^t, W) = \sigma(p_{t+h} - p_t | w^t, W)$ . Both measure uncertainty about price-level paths emanating forward from a given  $(w^t, W)$  pair.<sup>10)</sup>

The conditioning variables  $(w^t, W)$  are unknown, but our MCMC algorithm delivers a sample from their joint posterior distribution. By calculating  $\sigma(p_{t+h} - p_t | w^t, W)$  for every  $(w^t, W)$  pair in this sample, we obtain a sample from the posterior distribution  $p(\sigma(p_{t+h} - p_t) | y^T)$ , where  $y^T$  denotes the entire sample of observations and  $T$  is the length of the sample. We call the resulting values *smoothed conditional volatilities* because they are derived from a posterior for  $(w^t, W)$  that conditions on the full sample  $y^T$ .<sup>11)12)</sup>

10) Notice that the permanent and transient components both contribute terms that make this conditional variance increase with the forecast horizon  $h$ . The increment to the term involving volatility in the transient component is  $r_t \exp(h\sigma_r^2/2)$ , while the increment to the term involving volatility in the permanent component consists of the sum of  $q_t \exp(h\sigma_q^2/2)$  plus a number of other nonnegative terms. If the values associated with permanent and transient components were similar (i.e. if  $q_t \approx r_t$  and  $\sigma_q \approx \sigma_r$ ), the permanent component would grow more rapidly and dominate at long horizons. However, since we estimate  $q_t \ll r_t$  and  $\sigma_q < \sigma_r$ , the dominance of the permanent component is not a foregone conclusion, at least not at horizons that interest us.

11) We draw from the smoothed density  $p(w^t, W | y^T)$  rather than the filter densities  $p(w^t, W | y^t)$  because filter densities for  $t < 1948$  fail to exploit the identifying power of overlapping clean and noisy measures. Hence filter densities before and after World War II would be on a different footing vis a vis

Figure 8. Posteriors for Smoothed Conditional Volatilities 5 and 10 Years Ahead



Smoothed conditional volatilities for forecast horizons of 5 and 10 years are shown in figure 8. Solid lines again depict the posterior median at each date, and dotted lines portray the interquartile range. At least for these horizons, the evolution of price-level uncertainty closely resembles the movements in  $\sqrt{r_t}$  shown in figure 7. Price level uncertainty was greatest near the end of the Civil War, during the interwar period and two World Wars, and near the end of the sample. A smaller bump appears during the Great Inflation, but this again seems minor. The periods of lowest volatility occurred around the turn of the 20th century, in the late 1950s and early 1960s under Bretton Woods, and during the Great Moderation of the 1990s. The close resemblance between  $\sqrt{r_t}$  and smoothed conditional volatilities attests that transient bursts of inflation remain the dominant source of price-level uncertainty at horizons of 5 and 10 years. The dominance of transient variation follows from the fact that  $q_t \ll r_t$  and  $\sigma_q < \sigma_r$ .

Because periods of high and low volatility appear both before and after 1945, there is no unequivocal answer to whether there was greater price-level uncertainty before or after the war. On the contrary, the answer depends on the particular years being compared. We think it is sensible to focus on local peaks and

---

measurement error.

12) For  $t < T$ , a smoothed conditional volatility is not a posterior conditional-on- $y^t$  standard deviation, but we believe it is a better choice because the conditional-on- $y^t$  standard deviation would tell us more about measurement error than about price-level volatility.

Table 1. Peaks and Troughs in Smoothed Conditional Volatilities

Peaks	1864	1921	1980	2009
Troughs	1901	1963	1992	

Note: Median estimates,  $h = 10$  years.

troughs. As shown in table 1, median estimates of smoothed volatilities 10 years ahead peaked in 1864, 1921, 1980, and 2009 and reached troughs in 1901, 1963, and 1992. Dates of peaks and troughs for the 5-year horizon differ only slightly, so we concentrate on these 7 years.

Table 2 compares smoothed conditional volatilities in selected pre- and postwar years. The columns and rows refer, respectively, to particular pre- and postwar base years. Thus, the top left panel compares 1864 with 1963, while the bottom right compares 1921 with 2009. The top entry in each panel is the ratio of the postwar mean smoothed conditional standard deviation relative to that in the prewar base year. Entries in parentheses record the proportion of posterior sample paths on which conditional standard deviations are lower after the Second World War.

We begin by comparing peaks with peaks, viz. 1864 and 1921 v. 1980 and 2009. At postwar peaks, mean smoothed volatilities were approximately 40-75 percent lower than before the war, and the probability that uncertainty was greater at prewar peaks either approaches or exceeds 95 percent. Our calculations therefore suggest that price-level uncertainty was greater at prewar peaks. To some extent, this must reflect that pre-1945 crises were more traumatic than those occurring afterward (the Civil War, two World Wars, and the Great Depression versus the Great Inflation and Recession). But it might also reflect that policy authorities became more adept at crisis management.

By way of contrast, when comparing troughs with troughs (1901 v. 1963 and 1992), evidence favoring a postwar stabilization is weaker. Mean smoothed volatilities were roughly 40-50 percent lower after the Second World War, and there was less uncertainty on a preponderance of posterior sample paths (81-87 percent). But since uncertainty was higher after the war on a significant minority of paths (13-19 percent), the evidence is not entirely compelling.

One can of course also compare peaks with troughs. Not surprisingly, there was more uncertainty at prewar peaks than at postwar troughs. For instance, mean

Table 2. Relative Conditional Standard Deviations

	1864		1901		1921	
	5 years	10 years	5 years	10 years	5 years	10 years
1963	0.132	0.130	0.526	0.526	0.120	0.118
	(1.0)	(1.0)	(0.872)	(0.872)	(0.999)	(0.999)
1980	0.286	0.286	1.143	1.156	0.260	0.260
	(0.997)	(0.997)	(0.440)	(0.441)	(0.998)	(0.998)
1992	0.152	0.150	0.606	0.606	0.138	0.136
	(1.0)	(1.0)	(0.811)	(0.811)	(0.999)	(0.999)
2009	0.587	0.586	2.345	2.369	0.533	0.533
	(0.935)	(0.934)	(0.038)	(0.039)	(0.966)	(0.966)

Note: The top entry in each panel is the ratio of the postwar mean smoothed conditional standard deviations relative to that in the prewar base year. Entries in parentheses record the proportion of sample paths on which conditional standard deviations are lower in the postwar year.

smoothed volatilities for 1963 and 1992 are approximately 85 percent lower than those for 1864 and 1921, and the probability that price levels were more predictable at postwar troughs approaches 1. What to make of this comparison is unclear, however, for there is also evidence – albeit a bit weaker – that uncertainty was greater at postwar peaks than prewar troughs. For instance, mean smoothed volatilities in 2009 were more than double those in 1901, and the probability that uncertainty was greater during the financial crisis exceeds 95 percent.<sup>13)</sup>

In any case, no simple answer emerges to whether the price level was more predictable before or after 1945. Episodes of high and low volatility occur in both periods, precluding a blanket characterization. The trough-to-trough calculations confirm the promises of Ricardo and Keynes that a well-managed fiat regime can deliver outcomes comparable to or perhaps even better than those of a commodity standard.<sup>14)</sup> Absent big shocks such as wars, depressions, or financial panics, price level uncertainty can be moderated either by means of a link to gold, as under Bretton Woods or the classical gold standard, or via a well-managed fiat regime, as in the 1990s.

13) A comparison between 1901 and 1980 is inconclusive. Mean smoothed volatilities were about 15 percent higher in 1980, and conditional standard deviations were greater in 1980 on 55 percent of posterior sample paths.

14) See Sargent (2008) for accounts of their ideas and proposals.

## 5.4 Price-level Instability

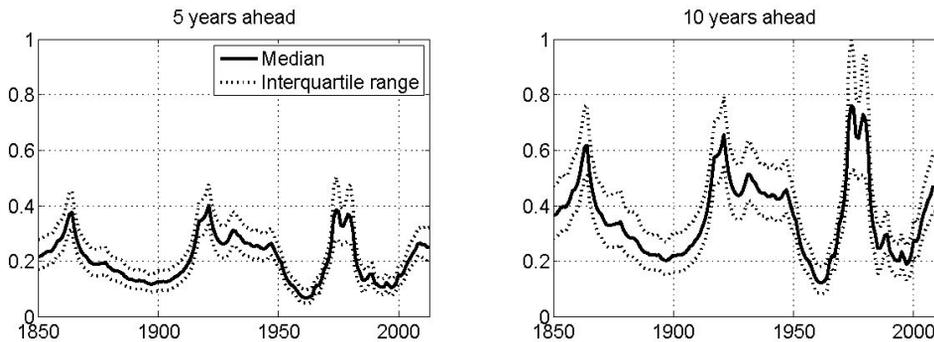
While conditional variances are suitable for measuring uncertainty, the concept of price stability seems different. We interpret ‘stability’ as describing *total* variation, not just *unpredictable* variation. For assessing price instability, we therefore compare second moments across dates. Our measure of instability is the square root of the conditional second moment (aka the conditional root mean square) for cumulative inflation,

$$crms(p_{t+h} - p_t | \omega^t, W) = \sqrt{\sigma^2(p_{t+h} - p_t | \omega^t, W) + h^2 \mu_t^2}. \quad (8)$$

This measures total price-level variation emanating forward from a given  $(\omega^t, W)$  pair.<sup>15)16)</sup>

Smoothed conditional root mean square statistics are shown in figure 9. Since the conditional second moment is the sum of the conditional variance plus the square of the conditional mean, the distinction between uncertainty and instability materializes only when the conditional mean  $h\mu_t$  is sufficiently far from zero. The

**Figure 9. Posteriors for Smoothed Conditional Root Mean Square Statistics 5 and 10 Years Ahead**



15) When central bankers speak of price stability, however, they sometimes seem to mean stability around an inflation target. This concept is closer to our notion of predictability. Readers who prefer this interpretation should therefore pay more attention to the results reported in section 5.3.

16) While the conditional variance for the log price level is the same as that for cumulative inflation, conditional second moments are not, because the conditional means differ. The conditional mean for the log price level  $h$  periods ahead is  $p_t + h\mu_t$ , so  $(p_t + h\mu_t)^2$  must be added to the conditional variance. Conditional second moments of the log price level therefore penalize both past and future cumulative inflation. Why past cumulative inflation should be penalized is not obvious, however, so we prefer to measure price instability by conditional second moments for cumulative inflation. Since variation in  $p_t$  dominates everything else, conditional second moments for the log price level increase with time and are greatest at the end of the sample.

only times this happened in our sample were the Great Inflation of the 1970s and, to a lesser extent, the financial crisis of 2009. The primary difference between the conditional root mean square statistics shown in figure 9 and the conditional standard deviations shown in figure 8 is that there are larger bumps in our instability measure in the 1970s. Indeed, median estimates of price instability in late 1970s are larger than prewar peaks. Otherwise, the evolution of conditional root mean square statistics resembles that of conditional standard deviations.

Among other things, this means that our trough-to-trough comparisons of uncertainty extend to our measure of instability. It follows that no compelling evidence emerges that the price level was more stable during the two postwar moderations than at the turn of the 20th century.

The main difference concerns peak-to-peak comparisons. For instance, while there was less price-level *uncertainty* in 1980 than at prewar peaks, there was about as much price-level *instability*. Mean smoothed *crms* statistics for 1980 are approximately 5 to 10 percent lower than those of 1864 and 1921 at the 5 year horizon and 5 to 10 percent higher at the 10 year horizon, and the odds that prices were more stable in one era over the other are not far from even (see the first and third columns of the 1980 row in table 3).

Comparing the prewar peaks with 2009 yields similar results. The interquartile range for  $\mu t$  was just 2-4 percent in 2009, but this is enough to push instability sta-

**Table 3. Relative Conditional Root-Mean-Square Statistics**

	1864		1901		1921	
	5 years	10 years	5 years	10 years	5 years	10 years
1963	0.198 (0.999)	0.218 (0.999)	0.581 (0.807)	0.595 (0.773)	0.187 (0.999)	0.208 (0.999)
1980	0.948 (0.553)	1.107 (0.434)	2.785 (0.025)	3.027 (0.027)	0.894 (0.603)	1.058 (0.469)
1992	0.285 (0.998)	0.323 (0.992)	0.839 (0.601)	0.884 (0.551)	0.269 (0.999)	0.309 (0.996)
2009	0.707 (0.834)	0.748 (0.773)	2.077 (0.061)	2.046 (0.085)	0.667 (0.880)	0.715 (0.823)

Note: The top entry in each panel is the ratio of the postwar mean smoothed *CRMS* relative to that in the prewar base year. Entries in parentheses record the proportion of sample paths on which *CRMS* statistics are lower in the postwar year.

tistics below the margin of significance. Relative to 1864 and 1921, ratios of mean smoothed *crms* statistics are approximately 0.67-0.75, but the probability that *crms* statistics were lower in 2009 falls below 0.9 (compare the 1864-2009 and 1921-2009 panels of tables 2 and 3). On a considerable minority of paths (12-23 percent), price levels 5 and 10 years ahead were less stable in 2009 than at prewar peaks.

Yet another difference concerns the comparison of the turn-of-the-20th-century moderation with the Great Inflation. While there were comparable levels of uncertainty in the two periods, the price level was much less stable in the latter (compare the 1901-1980 panel of tables 2 and 3). The average *crms* statistic for 1980 was almost 3 times greater than in 1901, and only 2.5 percent of posterior sample paths featured more stability in 1980.

### 5.5 Deflation Risk

Another aspect of price level dynamics that changed after the Second World War is the risk of deflation. To develop evidence on this feature of the data, we calculate smoothed deflation probabilities and compare them across dates.

As an analogue to our measures of uncertainty and instability, we want the probability of cumulative deflation  $p_{t+h} - p_t < 0$  going forward from a given  $(w^t, W)$  pair. Toward this end, let  $\xi_t^h$  be a sequence of potential future shocks  $\varepsilon_{\pi_s}, \varepsilon_{\mu_s}, \eta_{q_s}, \eta_{r_s}$ . According to the model,  $\xi_t^h$  has a normal unconditional distribution that we denote  $p_N(\xi_t^h)$ . Given a pair  $(w^t, W)$  along with a draw of potential future shocks  $\xi_t^h$ , a sequence of potential future price level realizations  $p_s(w^t, W, \xi_t^s)_{s=1}^h$  can be calculated by iterating on the state equations. We define an indicator variable

$$I_{deflation}(p_h(w^t, W, \xi_t^h), p_t(w^t, W)) = \begin{cases} 1 & \text{if } p_{t+h} < p_t, \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

that records whether cumulative inflation is positive or negative. Conditional on  $(w^t, W)$ , the probability of deflation is

$$dpr(w^t, W) \equiv pr(p_{t+h} - p_t < 0 | w^t, W), \quad (10)$$

$$= \int I_{deflation}(p_h(\omega^t, W, \xi_t^h), p_t(\omega^t, W)) p_N(\xi_t^h) d\xi_t^h.$$

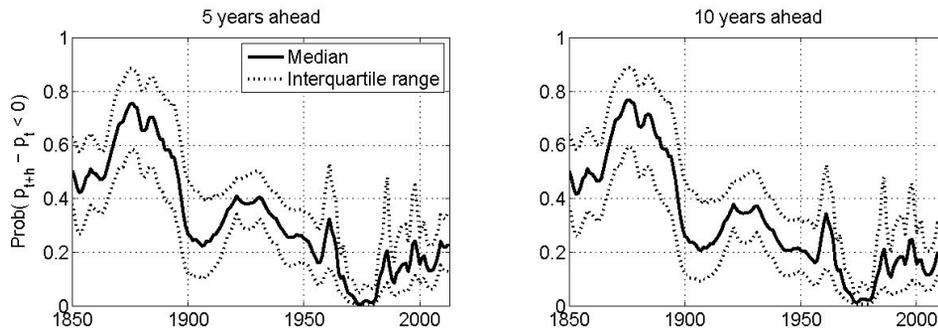
It is easy to approximate this probability by Monte Carlo integration. By calculating  $d_{pr}$  for every  $(\omega^t, W)$  pair in the posterior sample, we obtain its posterior distribution.

We call the resulting values *smoothed deflation probabilities* because they are derived from a posterior for  $\omega^t$  and  $W$  that conditions on the full sample  $y^T$ .

Figure 10 depicts the median and interquartile range for  $d_{pr}$  for forecast horizons of 5 and 10 years. Not surprisingly, low values of  $\mu_t$  and high values of  $\sigma(p_{t+h} - p_t)$  both contribute to high deflation risk. Thus, the probability of deflation was highest – above 70 percent – after the Civil War when  $\mu_t$  was negative and  $\sigma_t$  was high. The opposite scenario, periods with high  $\mu_t$  and low  $\sigma_t$ , did not occur in our sample, but deflation risk was lowest – below 5 percent – during the Great Inflation when high values of  $\mu_t$  trumped intermediate values of  $\sigma_t$ . The three price level moderations lie in between, with deflation probabilities on the order of 10-20 percent.<sup>17)</sup> Thus, low values of  $\mu_t$  trump low values of  $\sigma_t$  during moderations.

Table 4 compares selected years before and after the Second World War. Peaks in deflation risk were reached in 1876, 1921, and 1998, and troughs occurred in 1909 and 1979. We also include 2009 to measure deflation risk during the financial crisis. As before, the top row records the ratio of mean deflation risks before and after the

**Figure 10. Smoothed Deflation Probabilities 5 and 10 Years Ahead**



<sup>17)</sup> Because stochastic volatilities are represented as geometric random walks, our model generates very thick tails for future price levels even in tranquil times.

Table 4. Relative Deflation Probabilities

	1876		1909		1921	
	5 years	10 years	5 years	10 years	5 years	10 years
1979	0.082	0.071	0.198	0.181	0.136	0.125
	(0.995)	(0.995)	(0.887)	(0.885)	(0.993)	(0.992)
1998	0.396	0.381	0.958	0.971	0.656	0.671
	(0.918)	(0.922)	(0.532)	(0.518)	(0.746)	(0.714)
2009	0.368	0.320	0.889	0.814	0.609	0.562
	(0.951)	(0.954)	(0.528)	(0.553)	(0.840)	(0.828)

Note: The top entry in each panel is the ratio of the postwar mean probability of deflation relative to that in the prewar base year. Entries in parentheses record the proportion of sample paths on which the probability of deflation was lower in the postwar year.

war, and the bottom depicts the probability that  $dpr(w^t, \theta)$  was lower after the war.

At the 1998 postwar peak, mean deflation risk was roughly 35-60 percent lower than at the prewar peaks of 1921 and 1876 (see the first and third columns of the row corresponding to 1998). Conditional deflation probabilities were also lower than in 1876 on more than 90 percent of posterior sample paths. Relative to 1921, however, the evidence is less decisive:  $dpr(w^t, \theta)$  was higher in 1921 on 75 percent of posterior paths but was lower on the remaining 25 percent. Comparing 2009 with the prewar peaks yields similar results, with slightly stronger evidence of lower postwar deflation risk.

The postwar peak in  $dpr$  is comparable to the prewar trough (see the 1909-1998 panel), but the postwar trough is negligible in comparison with either of the prewar peaks (see the first and third columns of the 1979 row). A trough-to-trough comparison yields weak evidence of a postwar decline (see the 1909-1979 panel). At the 1979 postwar trough, mean deflation risk was approximately one-fifth as large as at the 1909 prewar trough, but the probability of decline fell short of 90 percent. Thus, on a nontrivial minority of posterior sample paths, there was less deflation risk at the prewar trough.

## VI. Discussion of Related Literature

Our paper is closely related to research on changing inflation persistence. For instance, Barsky (1987) documents that inflation was U.S. inflation was well approximated by a martingale-difference process before World War II and became strongly autocorrelated only during the Great Inflation. Benati (2008) corroborates these findings and shows more generally that inflation is weakly persistent whenever a monetary regime has a clearly defined nominal anchor. Our results and those of Stock and Watson (2007) are broadly consistent with their findings.

While our statistical model stays close to Stock and Watson's (2007) representation, a number of other papers extend it in various directions to capture other features of postwar US inflation. For instance, Cogley, et al. (2010) investigate autocorrelation in the inflation gap,  $\pi_t - \mu_t$ , and Cecchetti, et al. (2007) model log-volatility innovations as two-state Markov process, thus allowing jumps in  $r_t$  and  $q_t$ . Kozicki and Tinsley (2012) and Mertens (2011) introduce survey data on inflation forecasts to aid identification<sup>18)</sup>, Mertens (2011) and Stella and Stock (2012) develop multivariate forecasting models by incorporating data on other macroeconomic outcomes such as unemployment and the term structure of interest rates<sup>19)</sup> and Chan, et al. (2013) introduce a priori bounds on the random walk component  $\mu_t$ .<sup>20)</sup> Another worthy extension, especially for modeling 19th century inflation dynamics, would be to consider adding a stationary component to the log price level.

All of these extensions merit further research, but we chose not to undertake them here because we wanted to draw out the implications of a well-known and relatively simple statistical model for pre-World War II inflation volatility. We believe we have captured the main contours of the rise and fall in inflation volatility, and we encourage others to join us by studying how alternative model specifications affect the results.

---

18) Surveys do not extend far enough back in time to help us.

19) Coping with measurement error is more challenging in a multivariate environment, but this problem is well worth undertaking.

20) In the present context, estimates of  $\mu_t$  seem reasonable, making explicit bounds unnecessary. Whether bounds that are reasonable for an inflation-targeting regime would remain valid in the event of a crisis such as a great and dangerous war is open to doubt.

## VII. Concluding Remarks

Our analysis of US wholesale price inflation evinces recurring episodes of rising and falling price-level uncertainty and instability. Major outbreaks are associated with the Civil War, the two World Wars and Great Depression, and the Great Inflation and Great Recession. In each instance, a crisis disrupted pre-existing monetary arrangements and created considerable uncertainty about the future. In each case, policy makers eventually found a path back to price stability, but that took a long time: the average time from peak to trough was 30 years.

The notion that prices were more stable before or after the Second World War seems to rest on trough to peak comparisons. When comparing peaks with peaks and troughs with troughs, no compelling evidence emerges that the price level was more stable in one era than the other. There was less uncertainty at peaks after the Second World War, but this might be due to the fact that postwar crises were less severe.

Evidence emerges for three 'great price-level moderations,' one near the turn of 20th century, another under Bretton Woods, and a third in the 1990s. Well-managed systems sometimes involved a link to gold – as under the classical gold standard and Bretton Woods – but the experience of the 1990s proves that a well-managed fiat regime can achieve the same end, as Ricardo and Keynes promised. Thus a link to gold or some other commodity is unnecessary for moderating price level volatility.

Nor does a commodity standard provide a guarantee of permanent price stability. The major challenge is coping with big shocks such as wars, depressions, and financial panics. Neither a commodity nor a fiat regime is immune to big shocks, and convertibility has typically been suspended in times of crisis. During the US Civil War, policy makers in both the North and the South stopped backing their currencies with gold or silver because issues of war finance became their paramount concerns. Convertibility was restored in the North only in 1879, and never in the South. The classical gold standard and Bretton Woods system ended not because the authorities thought they had discovered better methods for maintaining price stability but because other economic objectives supervened. Because main-

taining convertibility during a crisis was unlikely, there never was a believable unconditional commitment to a commodity standard. Before the breakdown of Bretton Woods we saw mixed standards, with convertibility during tranquil times and suspensions and reversions to fiat money after big shocks.

Our analysis is cast in terms of a purely *statistical* model, so we have nothing formal to say about why someone might care about the uncertainty and/or the stability of inflation and how they have changed over time. But plenty of economic models do suggest that we should care. For example, models in the tradition of Lucas (1973) make social welfare depend on the uncertainty of the log price level, but not its stability. Other models in the tradition of Lucas (2000) focus on welfare effects of the stability of inflation, not its uncertainty. In other more general macroeconomic models, people might care about both price level stability and price level uncertainty. Our statistical characterizations of price level stability and price level uncertainty set the stage for a quantitative economic analysis of their welfare consequences.

## VIII. References

- Barsky, R.B. (1987), "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics*, Vol. 19, 324.
- Benati, L. (2008), "Investigating Inflation Persistence Across Monetary Regimes," *Quarterly Journal of Economics*, Vol. 123, pp. 1005-1060.
- Cecchetti, S.G., P. Hooper, B.C. Kasman, K.L. Schoenholtz, and M.W. Watson (2007), "Understanding the Evolving Inflation Process," *U.S. Monetary Policy Report* No.1.
- Chan, J.C.C., G. Koop, and S. Potter (2013), "A New Model of Trend Inflation," *Journal of Business and Economic Statistics*, Vol. 34, pp. 91-106.
- Cogley, T. and T.J. Sargent (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post War U.S.," *Review of Economic Dynamics*, Vol. 8, pp. 262-302.
- Cogley, T., G.E. Primiceri, and T.J. Sargent (2010), "Inflation-Gap Persistence in the U.S.," *American Economic Journal: Macroeconomics*, Vol. 2, pp. 43-69.
- Cogley, T., T.J. Sargent, and P. Surico (2012), "Price-Level Uncertainty and Stability in the U.K.," unpublished manuscript, New York University and London Business School.
- Hanes, C. (1998), "Consistent Wholesale Price Series for the United States, 1860-1990," in Trevor J. O. Dick, editor, *Business Cycles Since 1820: New International Perspectives from Historical Evidence*, Edward Elgar.
- Hanes, C. (2006), *Prices, Historical Statistics of the United States: Earliest Times to the Present*, Millennial Edition, Cambridge University Press.
- Jacquier, E., N.G. Polson, and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, Vol. 12(4), pp. 371-89.
- Kozicki, S. and P.A. Tinsley (2012), "Effective Use of Survey Information in Estimating the Evolution of Expected Inflation," *Journal of Money, Credit, and Banking*, Vol. 44, pp. 145-169.
- Lucas, R.E., Jr. (1973), "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, Vol. 63(3), pp. 326-34.

- Lucas, R.E., Jr. (2000), "Inflation and Welfare," *Econometrica*, Vol. 68(2), pp. 247-274.
- Mertens, E. (2011), "Measuring the Level and Uncertainty of Trend Inflation," Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series, Paper 2011-42.
- Romer, C.D. (1986a), "Is the Stabilization of the Postwar Economy a Figment of the Data?" *American Economic Review*, Vol. 76, pp. 314-334.
- Romer, C.D. (1986b), "Spurious Volatility in Historical Unemployment Data," *Journal of Political Economy*, Vol. 94, pp. 1-37.
- Sargent, T.J. (2008), "Evolution and Intelligent Design," *American Economic Review*, Vol. 98(1), pp. 5-37.
- Shephard, N. (2013), "Martingale Unobserved Components Models," unpublished manuscript, University of Oxford.
- Stella, A., and J.H. Stock (2012), "A State-Dependent Model for Inflation Forecasting," Board of Governors of the Federal Reserve System, International Finance Discussion Paper No. 1062.
- Stock, J.H. and M.W. Watson (2007), "Why Has Inflation Become Harder to Forecast?" *Journal of Money Credit and Banking*, Vol. 39(1), pp. 3-33.
- Warren, G.F., and F.A. Pearson (1933), "*Prices*," Wiley.

## Appendices

### A A Markov Chain Monte Carlo algorithm for simulating the posterior

We use a Metropolis-within-Gibbs algorithm to simulate the posterior distribution for model parameters and latent states. Let  $S^T = (\pi^T, \mu^T, m^T)$  and  $v^t = (r^T, s^T)$  represent histories of states through date  $T$ . Then the parameter blocks are

- $v_t | v_{\neq t}, S^T, W, \rho_m, \sigma_m$
- $W | v^T, S^T, \rho_m, \sigma_m$
- $S^T = | v^T, W, \rho_m, \sigma_m$
- $\rho_m, \sigma_m | v^T, S^T, W$ .

A Metropolis-within-Gibbs algorithms requires the conditional posterior density or kernel for each of these blocks.

#### A.1 $p(v_t | v_{\neq t}, S^T, W, \rho_m, \sigma_m, y^T)$

To sample the volatilities  $v_t = (r_t, s_t)$ , we adopt a bivariate version of the Jacquier, et al. (1994) single-move sampler. Conditional on the latent states  $S^T = (\pi^T, \mu^T, m^T)$ , the innovations

$$\zeta_t = \begin{bmatrix} \pi_t - \mu_t \\ \mu_t - \mu_{t-1} \end{bmatrix} \quad (\text{A-1})$$

are observable. Given our assumptions,  $\zeta_t$  is conditionally normal with mean zero and variance

$$H_t = \begin{bmatrix} r_t & 0 \\ 0 & q_t \end{bmatrix} \quad (\text{A-2})$$

We want the conditional posterior  $p(v_t | v_{\neq t}, \zeta^T, W, \rho_m, \sigma_m, y^T)$ , where  $v_{\neq t}$  represents the history  $v^T$  with  $v_t$  removed,  $v_{\neq t} = v_1, \dots, v_{t-1}, v_{t+1}, \dots, v_T$ . Because measurement errors are independent of other innovations in the model, knowl-

edge of  $y^T$  and  $m^T$  are redundant given knowledge of  $\pi^T$  and  $\mu^T$ , as is knowledge of  $\rho_m$  and  $\sigma_m$ . The problem therefore reduces to sampling from the conditional posterior  $p(v_t | v_{\neq t}, \zeta^T, W)$ .

According to Bayes' theorem, this conditional posterior density is

$$p(v_t | v_{\neq t}, \zeta^T, W) \propto p(\zeta^T | v^T, W) \cdot p(v_t | v_{\neq t}, W). \quad (\text{A-3})$$

Using the prediction error decomposition, the conditional likelihood function can be expressed as

$$\begin{aligned} p(\zeta^T | v^T, W) &= \prod_{j=1}^T p(\zeta_j | \zeta^{j-1}, v^T, W), \\ &= \prod_{j=1}^T p(\zeta_j | v_j), \end{aligned} \quad (\text{A-4})$$

where the second equality follows from the fact that the conditional densities on the right hand side are gaussian and that  $(\zeta^{j-1}, v_{\neq j}, W)$  affect neither the conditional mean nor the conditional variance for  $\zeta_j$ . Furthermore, because  $v_t$  is Markovian, the only elements of  $v_{\neq t}$  that are relevant for the conditional prior are for  $v_t$  are  $v_{t-1}$  and  $v_{t+1}$ .<sup>21)</sup> Hence the conditional posterior simplifies to

$$p(v_t | \zeta_t, v_{t-1}, v_{t+1}, W) \propto p(\zeta_t | v_t) \cdot p(v_t | v_{t-1}, v_{t+1}, W). \quad (\text{A-5})$$

As in Cogley and Sargent (2005), we adopt a Metropolis-Hastings step with proposal density  $f(v_t) = p(v_t | v_{t-1}, v_{t+1}, W)$ , thus drawing a trial value for  $\ln v_t$  from a normal density with mean  $\mu_t = (\ln v_{t+1} + \ln v_{t-1})/2$  and variance  $\Omega = (1/2)W$ .

The acceptance probability for date  $t$ , scan  $k$  is  $\min(1, a_{kt})$ , where

$$\begin{aligned} \alpha_{kt} &= \frac{p(\zeta_t | v_t^*) p(v_t^* | v_{t-1}, v_{t+1}, W)}{p(\zeta_t | v_t^{k-1}) p(v_t^{k-1} | v_{t-1}, v_{t+1}, W)} \frac{f(v_t^{k-1})}{f(v_t^*)} \\ &= \frac{|H_t^*|^{-1/2} \exp\{-0.5 \zeta_t' (H_t^*)^{-1} \zeta_t\}}{|H_t^{k-1}|^{-1/2} \exp\{-0.5 \zeta_t' (H_t^{k-1})^{-1} \zeta_t\}}. \end{aligned} \quad (\text{A-6})$$

21) For the end points  $t = 1, T$ , the details differ slightly because we can condition only on one adjacent realization.

If the proposed value  $v_t^*$  is accepted, the chain advances to  $v_t^k = v_t^*$ . Otherwise,  $v_t^k = v_t^{k-1}$ .

### A.2 $p(W|v^T, S^T, \rho_m, \sigma_m, y^T)$

Given  $v^T$ , the log-volatility innovation  $W^{1/2}(\eta_{rt}, \eta_{qt})'$  are observable. All other conditioning information is irrelevant because  $\eta_{rt}$  and  $\eta_{qt}$  are independent of all other shocks in the model. The conditional posterior therefore reduces to  $p(W|v^T)$ . The log-volatility innovations are iid normal with mean zero and variance  $W$ . Since the prior on  $W$  is inverse-Wishart and the likelihood function is gaussian, the posterior is also  $1/W$ . Hence we sample  $W$  by drawing from an  $1/W$  density.

### A.3 $p(S^T|v^T, W, \rho_m, \sigma_m, y^T)$

Conditional on histories for  $y^T$  and  $v^T$  and the parameters  $\rho_m, \sigma_m$ , and  $W$ , equations (1)-(4) form a linear-Gaussian state-space system with known transition arrays and innovation variances. It follows that the posterior conditional density for the history of states  $S^T$  is Gaussian. To simplify computations, we introduce a tiny bit of measurement error for  $y_t$  whose variance equals machine epsilon times  $I_{\dim(y_t)}$ . With this simplification,  $S^T$  can be sampled via Carter and Kohn's (1994) forward-filter, backward sampler.<sup>22)</sup>

### A.4 $p(\rho_m, \sigma_m|v^T, S^T, W, y^T)$

While analogs to blocks 1-3 appear in Cogley and Sargent (2005), block 4 is new. The model is designed, however, so that inference about measurement error parameters is straightforward. Because measurement innovations  $\varepsilon_{mt}$  are independent of all other shocks in the model, all conditioning information other than  $m^T$  is redundant. The conditional posterior therefore reduces to  $p(\rho_m, \sigma_m|m^T)$ , the parameters of an autoregression for and observed series  $m^T$ . We adopt a normal-IG prior on  $(\rho_m, \sigma_m)$ . Because the measurement innovations are iid normal,

22) Carter and Kohn's algorithm can be modified to handle zero measurement error, but we felt that an infinitesimal improvement in accuracy did not warrant the extra complication.

the conditional posterior is also normal-IG. We sample  $(\rho_m, \sigma_m)$  from this posterior.

### A.5 Convergence diagnostics

The Markov chain was run for 4 million steps, and 2 million steps were discarded as burn-in. Convergence was diagnosed by inspecting recursive mean plots and the outcomes of parallel chains with different initial conditions. The output of alternative chains differed in no interesting respects.

### A.6 Additional references

Carter, C.K. and R. Kohn, 1994, On Gibbs Sampling for State-Space Models, *Biometrika* 81, 541-55.3.

Jacquier, E., N.G. Polson, and P.E. Rossi, 1994. Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, vol. 12(4), 371-89.

## B Deriving the conditional mean and variance of cumulative inflation

The log price level can be represented as

$$p_{t+1} = p_t + \mu_t + \sqrt{q_{t+1}} \varepsilon_{\mu t+1} + \sqrt{r_{t+1}} \varepsilon_{\pi t+1} \quad (\text{A-7})$$

The conditioning set at  $t$  is denoted  $(w^t, W)$  where  $w^t = (\pi^t, \mu^t, r^t, s^t)$  represents the history of the time-varying states and  $W$  is the covariance matrix for log-volatility innovations  $(\eta_{rt}, \eta_{qt})'$ . For a one-step ahead forecast, the conditional mean is

$$E(p_{t+1} | w^t, W) = p_t + \mu_t, \quad (\text{A-8})$$

and the prediction error is

$$p_{t+1} - E(p_{t+1} | w^t, W) = \sqrt{q_{t+1}} \varepsilon_{\mu t+1} + \sqrt{r_{t+1}} \varepsilon_{\mu t+1} \quad (\text{A-9})$$

Because the state innovations are standard normal and independent of log volatilities, the conditional variance is

$$\text{var}(p_{t+1} | w^t, W) = E(q_{t+1} + r_{t+1} | w^t, W). \quad (\text{A-10})$$

The variables  $r_{t+1}$  and  $q_{t+1}$  are conditionally lognormal one step ahead with means  $r_t \cdot \exp(\sigma_r^2/2)$  and  $q_t \cdot \exp(\sigma_q^2/2)$ , respectively. It follows that

$$\text{var}(p_{t+1}|\omega^t, \theta) = q_t \cdot \exp(\sigma_q^2/2) + r_t \cdot \exp(\sigma_r^2/2). \quad (\text{A-11})$$

Now consider a two-step ahead forecast. Since the log price level can be expressed as

$$p_{t+2} = p_{t+1} + \mu_{t+2} + \sqrt{r_{t+2}} \varepsilon_{\pi t+2} \quad (\text{A-12})$$

the conditional mean is

$$E(p_{t+2}|\omega^t, W) = E(p_{t+1}|\omega^t, W) + E(\mu_{t+2}|\omega^t, W) = p_t + 2\mu_t, \quad (\text{A-13})$$

and the prediction error is

$$\begin{aligned} p_{t+2} - E(p_{t+2}|\omega^t, W) &= (p_{t+1} - E(p_{t+1}|\omega^t, W)) + (\mu_{t+2} - E(\mu_{t+2}|\omega^t, W)) \quad (\text{A-14}) \\ &+ \sqrt{r_{t+2}} \varepsilon_{\pi t+2} = 2\sqrt{q_{t+1}} \varepsilon_{\mu t+1} + \sqrt{q_{t+2}} \varepsilon_{\mu t+2} + \sqrt{r_{t+1}} \varepsilon_{\mu t+1} + \sqrt{r_{t+2}} \varepsilon_{\pi t+2}. \end{aligned}$$

Hence the two-step-ahead conditional variance is

$$\begin{aligned} \text{var}(p_{t+2}|\omega^t, W) &= 4q_t \cdot \exp(\sigma_q^2/2) + r_t \cdot \exp(\sigma_r^2/2) \quad (\text{A-15}) \\ &+ E(\sqrt{q_{t+2}} \varepsilon_{\mu t+2}^2 + \sqrt{r_{t+2}} \varepsilon_{\pi t+2}^2 | \omega^t, W). \end{aligned}$$

By the law of iterated expectations, the last two right-hand terms can be expressed as

$$E(\sqrt{x_{t+2}} \varepsilon_{x t+2}^2 | \omega^t, W) = E[E(\sqrt{x_{t+2}} \varepsilon_{x t+2}^2 | \omega^{t+1}, W) | \omega^t, W] \quad (\text{A-16})$$

$x = q, r$ . The inner expectation is a one-step conditional variance,

$$E(\sqrt{x_{t+2}} \varepsilon_{x t+2}^2 | \omega^{t+1}, W) = x_{t+1} \exp(\sigma_x^2/2), \quad (\text{A-17})$$

and the outer expectation is a one-step-ahead forecast of that variance,

$$\begin{aligned} E[x_{t+1} \exp(\sigma_x^2/2) | \omega^t, W] &= E(x_{t+1} | \omega^t, W) \exp(\sigma_x^2/2), \quad (\text{A-18}) \\ &= x_t \exp(\sigma_x^2/2) \exp(\sigma_x^2/2), \\ &= x_t \exp(2\sigma_x^2/2) \end{aligned}$$

It follows that the two-step ahead conditional variance is

$$\begin{aligned} \text{var}(p_{t+2}|\omega^t, W) &= q_t [4\exp(\sigma_q^2/2) + \exp(2\sigma_q^2/2)] + r_t [\exp(\sigma_r^2/2) + \exp(2\sigma_r^2/2)], \\ &= q_t \sum_{j=1}^2 (2-j+1)^2 \exp(j\sigma_q^2/2) + r_t \sum_{j=1}^2 \exp(j\sigma_r^2/2). \quad (\text{A-19}) \end{aligned}$$

Now set the forecast horizon  $h = 3$ . Since the log price level can be expressed as

$$p_{t+3} = p_{t+2} + \mu_{t+3} + \sqrt{r_{t+2}} \varepsilon_{\pi t+3}, \quad (\text{A-20})$$

the conditional mean is

$$E(p_{t+3}|\omega^t, W) = E(p_{t+2}|\omega^t, W) + E(\mu_{t+3}|\omega^t, W) = p_t + 3\mu_t, \quad (\text{A-21})$$

and the prediction error is

$$\begin{aligned} p_{t+3} - E(p_{t+3}|\omega^t, W) &= (p_{t+2} - E(p_{t+2}|\omega^t, W)) + (\mu_{t+3} - E(\mu_{t+3}|\omega^t, W)) + \sqrt{r_{t+2}} \varepsilon_{\pi t+3}, \quad (\text{A-22}) \\ &= 3\sqrt{q_{t+1}} \varepsilon_{\mu t+1} + 2\sqrt{q_{t+2}} \varepsilon_{\mu t+2} + \sqrt{q_{t+3}} \varepsilon_{\mu t+3} + \sqrt{r_{t+1}} \varepsilon_{\pi t+1} + \sqrt{r_{t+2}} \varepsilon_{\pi t+2} + \sqrt{r_{t+3}} \varepsilon_{\pi t+3}, \end{aligned}$$

By imitating the argument for a two-step ahead conditional variance, one can demonstrate that the three-step ahead conditional variance is

$$\text{var}(p_{t+3}|\omega^t, W) = q_t \sum_{j=1}^3 (3-j+1)^2 \exp(j\sigma_q^2/2) + r_t \sum_{j=1}^3 \exp(j\sigma_r^2/2). \quad (\text{A-23})$$

By continuing the recursion, one can show that the  $h$ -step ahead conditional mean and variance for cumulative inflation are

$$E(p_{t+h} - p_t | \omega^t, W) = h\mu_t, \quad (\text{A-24})$$

$$\text{var}(p_{t+h} - p_t | \omega^t, W) = q_t \sum_{j=1}^h (h-j+1)^2 \exp(j\sigma_q^2/2) + r_t \sum_{j=1}^h \exp(j\sigma_r^2/2),$$

respectively.