THREE ESSAYS ON INFLATION EXPECTATIONS IN EMPIRICAL MACROECONOMICS

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
doctor rerum politicarum
(Dr. rer. pol.)

am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Max Conrad Diegel, M.Sc.
aus Bergisch-Gladbach

Berlin, Oktober 2022

Acknowledgements

Berlin, Oktober 2022   *Max Diegel*

Declaration of Co–Authorship

This dissertation consists of three (working) papers. One paper was written in collaboration with one co-author. The contribution in conception, implementation and drafting for the three chapters can be summarized as follows:

1. Long-term Inflation Expectations and the Transmission of Monetary Policy Shocks: Evidence from a SVAR analysis *with Dieter Nautz*
   Contribution: 50%

2. The time-varying Credibility of the Fed's Inflation Target in a Model with Asymmetric Information
   Contribution: 100%

3. Structural Vector Autoregressions with Common Factor Stochastic Volatility
   Contribution: 100%

Contents

Contents

List of Figures

List of Tables

Introduction and Overview

Inflation expectations have become a major indicator to assess the credibility of a central bank's inflation target. Therefore, the monetary policy and the communication strategy of inflation targeting central banks like the Federal Reserve Bank of the US (Fed) and the European Central Bank (ECB) is concerned with a firm anchoring of inflation expectations at their targets. Yet, long-term inflation expectations did not always coincide with the inflation targets in the recent past. In the aftermath of the 2008 Financial Crisis, inflation expectations in the US and Europe fell short of their respective central banks' inflation targets. Consequently, unconventional monetary policy measures by the Fed and the ECB sought to re-anchor too low inflation expectations back to the target level. In the 2020s, the challenge faced by central banks is completely reversed: International supply chain disruptions that started to show in mid-2021 as a result of the Covid-19 pandemic and the escalation of the conflict between Russia and Western Europe in 2022 have led to a surge of inflation to levels and an increase in macroeconomic volatility not seen since the beginning of the Great Moderation. The adoption of average inflation targeting by the Fed (see Powell 2020) and the ECB (2021), alongside the slow shift to the challenges of this new regime may have given an impression of higher tolerance for inflation, see e.g. Schnabel (2022). In the course of 2022, longer-term inflation expectations, even those of professional forecasters, have steadily increased to levels well above the inflation targets. This poses a potential threat to the credibility of central bank's inflation targets and raises the risks of a de-anchoring of inflation expectations.

This dissertation examines the anchoring of inflation expectations from three perspectives. Parts of the following work were conducted within the DFG-project "The Anchoring of Inflation Expectations." While the papers were motivated by the too-low inflation experience, the recent rise in inflation has brought inflation expectations back into the limelight of current discussions among central bankers and macroeconomists for yet another reason. The

findings in this dissertation generally apply when inflation expectations are above target.[1]

The first chapter, which is joint work with Dieter Nautz, investigates the role of long- term inflation expectations for the monetary transmission mechanism and the conduct of monetary policy in a structural VAR framework. It is motivated by preceding work of Hachula and Nautz (2018) and Nautz et al. (2019), who find that long-term inflation expectations react to a macro news shock in bi-variate models of short- and long-term inflation expectations. The macro-news shock identified in bi-variate SVARs is a conglomerate of all structural economic shocks related to short-term macroeconomic developments. Yet, by abstracting from economic key variables, like economic activity, inflation, and interest rates, bi-variate models of short- and long-term inflation expectations cannot account for the various shocks considered by macroeconomic theory. As a result, following Leduc et al. (2007) an ever-growing literature embeds inflation expectations data into empirical models.[2] In line with a re-anchoring channel of monetary policy, Andrade et al. (2016) and Doh and Oksol (2018) find that the central banks' announcements of asset purchase programs play an active role in steering inflation expectations toward the inflation target. However, the literature does not show how the management of inflation expectations by the central bank affects the dynamics of actual inflation and the transmission of monetary policy shocks.

To fill this gap, we empirically investigate the role of long-term inflation expectations for the monetary transmission mechanism. Following D'Amico and King (2017) and Jarociński and Karadi (2020), we employ a minimal set of uncontroversial sign restrictions to identify a monetary policy shock in a structural VAR including inflation expectations. In contrast to Clark and Davig (2011), our results show that long-term inflation expectations do respond to monetary policy shocks. On impact, monetary policy shocks account for a non-negligible fraction of the variation of long-term inflation expectations.

To shed more light on the role of long-term inflation expectations for the monetary transmission mechanism, we compute two counterfactual scenarios. The first scenario shows that, in line with a re-anchoring channel, the immediate response long-term inflation expectations contributes to the reaction of actual inflation to a monetary policy shock. The second scenario shows that the endogenous reaction of monetary policy to expectations shocks has contributed to stabilizing expectations during the zero lower bound period. These results suggest that monetary policy can act to re-anchor inflation expectations if

---

[1]This dissertation does not explicitly account for non-linearities that may arise from inflation being below or above target.

[2]See for example Canova and Gambetti (2010), Del Negro and Eusepi (2011), Fuhrer (2012),and Coibion et al. (2018).

required.

The second chapter uses a macro model with asymmetric information, unobserved compo-
nents, and an explicit learning mechanism to strengthen the theoretical link with macroeco-
nomic theory, compared to the SVAR approach of the first chapter. Since the credibility of
a central bank's inflation target should not be assumed *a priori* to be always perfect, even
in the presence of officially announced targets, a key feature of the model is that it allows
for an explicit notion time-varying credibility and anchoring of the perceived target by the
public to the actual inflation target. This model contributes to the re-emerging literature
that assumes imperfectly informed agents learn about the economy and form expectations
based on recent data.

    While most papers focus on the learning from short-term inflation surprises (see e.g.
Carvalho et al. 2022; Jorgensen and Lansing 2022), Chapter two models how the public
learns from central bank's interest rate policy about the imperfectly observed and potentially
time-varying inflation target. To that end, the learning mechanism of Kozicki and Tinsley
(2005) is extended by allowing for breaks in the learning rate and the shock variances
according to the broad US monetary policy regimes. Implications for the optimal rate of
learning and the degree of anchoring are derived. The breaks in variances and the learning
rate give rise to time-variation, including in the degree of anchoring. To estimate the model
efficiently in a Bayesian framework, I extend the precision based framework for unobserved
components models of Chan and Eisenstat (2018) to the multivariate case.

    The estimated model shows that imperfect credibility is most pronounced after the
Volcker Disinflation and to a lesser extent in the aftermath of the Great Financial Crisis.
While credibility was poor due a too-high perceived inflation target by the public in the
former, the reverse is true for the latter period. The estimated learning rates are lower than
the optimal rate in all monetary policy regimes.

The third chapter focuses on yet another dimension of the anchoring of inflation expectations,
namely volatility. In particular, the chapter investigates whether there are common factors
in the volatilities of macro variables and US inflation expectations. However, the vast
majority of the SVAR literature that allows for stochastic volatility (SV) in the shocks
assumes that the volatilities evolve *independently* and, thus, ignores possible common sources
of the time-variation in volatilities. Therefore, I suggest a new model that allows for a
flexible number of common factors in the time-varying volatilities. This model nests popular
alternatives like the common stochastic volatility (CSV) model of Carriero et al. (2016),

which allows for just one factor that scales the entire residuals co-variance matrix of a reduced form VAR. To distinguish the new model from existing models, I refer to it as the Common Factor Stochastic volatility (CFSV) model. The model is based on a non-centered parametrization that allows to check the hypothesis of a reduced rank in the stochastic volatilities and linear restrictions on the loadings of the volatility factors via Savage Dickey Density Ratios (SDDR), thus avoiding the computation of the marginal likelihood that is particularly costly in non-linear models. Simulation evidence demonstrates that the model correctly recovers the reduced-rank volatility structure in a recursively identified SVAR.

Finally, I use the CFSV model to revisit the application of Clark and Davig (2011), who estimate an SVAR with a set of typical US macro variables and short- and long-term inflation expectations while allowing independent stochastic volatility (SV). Allowing for CFSV reveals that only two factors drive the time-variation of the five shocks. While the shock to long-term inflation expectations has its own volatility factor, the second factor loads onto both actual inflation and short-term inflation expectations. Such volatility spill-overs suggest that analysis based on independent SV might underestimate the adverse effect of increased inflation volatility that occurs via short-term inflation expectations. Thereby, the range of inflation outcomes might be underestimated as well. This could be particularly important in the environment of increased inflation volatility following the Covid-19 pandemic.

Chapter 1

Long-term Inflation Expectations and the Transmission of Monetary Policy Shocks: Evidence from a SVAR analysis

with Dieter Nautz

For copyright reasons, this chapter (p. 2 - 40) is not included in the online version of the dissertation. An electronic version of the article can be accessed at https://doi.org/10.1016/j.jedc.2021.104192

Chapter 2

Time-varying Credibility, Anchoring and the Fed's Inflation Target

"A fuller understanding of the public's learning rules would improve the central bank's capacity to assess its own credibility, to evaluate the implications of its policy decisions and communications strategy." Bernanke (2007)

2.1 Introduction

Central banks, like the Federal Reserve Bank (Fed), interpret price stability as inflation rates being close to an implicit or explicit inflation target. Thus, to achieve their inflation target, a central bank should ideally convey that it is credible in fulfilling its mandate of price stability. However, credibility is fragile and, once lost, can be costly to re-establish, as argued forcefully by Goodfriend (2004) for example. Therefore, monitoring the state of credibility at all times is indispensable.

Over most of the postwar period, the exact inflation target of the Fed was unknown to the public and probably time-varying. Therefore, the inflation target perceived by the public may sometimes deviate from the actual inflation target, creating a *credibility gap*. To help anchor public perceptions at, and strengthen the credibility of its target, the Fed has changed its communication strategy significantly since the early 1960s. For example, to foster understanding of the policy actions taken to reduce the high inflation rates of the late 1970s, "on October 6, 1979, the Fed broke sharply with its tradition of saying little in public about its actions" (Goodfriend 2007, p. 51) and explained them to a wider public. In January 2012, the Federal Open Market Committee (FOMC) announced the numerical

inflation target of 2 percent for the annual inflation rate of the Personal Consumption Expenditure (PCE) index. Later, it redefined the target to be an *average* inflation target. Hence, even official target announcements can be subject to change. Moreover, they do not necessarily eliminate asymmetric information and imperfect credibility. For example, Coibion et al. (2020) find that, in the US, 60% of survey respondents among firms and 40% among households said they do not know what the Fed's inflation target was. Therefore, the target perceived by the public need not automatically coincide with the officially announced inflation target by the central bank. Modeling the Fed's inflation target as time-varying and not directly observable by the public may still be an accurate representation of the Fed's monetary policy, even after 2012 and especially for the purpose of gauging potentially imperfect credibility.[20]

As pointed out in the lead quote by Bernanke (2007), knowledge of the public learning mechanism is important for assessing credibility and the impact of different communication strategies. Therefore, the aim of this paper is to contribute to a better understanding of the time-varying credibility of the Fed's inflation target by estimating how the public learns about the inflation target from the Fed's monetary policy.

This paper. The starting point of this paper is the learning mechanism of Kozicki and Tinsley (2005). In their model, the monetary policy rate is set by a Taylor rule with an unobservable and time-varying inflation target $\pi_t^\star$. Due to asymmetric information, the public has to solve signal extraction problem to learn about the Fed's inflation target. As a result, it updates the perceived target π_t^P in response to the observed interest rate policy of the central bank. Since $\pi_t^\star$ and π_t^P need not coincide, the model allows for a precise notion of imperfect credibility, the credibility gap. This paper then extends the analysis of Kozicki and Tinsley (2005) along three important dimensions.

First, I derive the learning rate that is optimal from the central bank's perspective, thus providing an important benchmark for assessing empirical estimates of the learning rate. To that end, I propose the expected squared credibility gap at medium- to long horizons as an indicator of the degree of de-anchoring of π_t^P from $\pi_t^\star$. This de-anchoring indicator naturally summarizes two key aspects of anchoring: Persistence and variance of the credibility gap. In the spirit of Jorgensen and Lansing (2022), the optimal learning rate is then found by minimizing this de-anchoring indicator. However, in contrast to the literature where

---

[20]Despite the official inflation target, Shapiro and Wilson (2019) find that the Fed more likely has targeted a rate of inflation that is slightly lower than the announced 2%, even after 2012. Moreover, average inflation targeting implies a certain degree of time-variation in the inflation target because the central may temporarily aim for inflation above or below 2%.

agents learn from inflation surprises instead of monetary policy, a learning rate of zero is generally not optimal because it implies that the perceived target is unrelated to the inflation target.[21] Therefore, the derived optimal rate adds an interesting new perspective on the relation between learning and anchoring. As a result of the non-zero optimal learning rate, a response of the perceived target to monetary policy can be desirable because it helps to anchor public long-term inflation beliefs at the actual inflation target, as suggested also by the SVAR evidence in Diegel and Nautz (2021).

Second, I allow for breaks in the structural shock variances and the learning rate across US monetary policy regimes. Kozicki and Tinsley restrict the shock variances and the learning rate to be constant. However, the literature demonstrates that changes in shock variances are important to fit the evolution of the Fed's monetary policy framework through the postwar period; see e.g. Primiceri (2005), Sims and Zha (2006b), and Belongia and Ireland (2016). The dates of the regime changes are motivated by narrative historical evidence of Goodfriend (2004, 2007), the 2012 announcement of the 2% target, and findings of the structural VAR literature; see e.g. Brunnermeier et al. (2021). Since the optimal learning rate depends on the shock variances, different monetary regimes may also imply different optimal learning rates. To capture the effect of the changes to the Fed's communication strategy on agent's learning behavior, I allow the learning rate to vary across regimes as well. To the best of my knowledge, this is the first paper to estimate a non-constant learning rate that links public perceptions directly to the inflation target. This allows to check whether the announcement has had the desired effect: In that case, it should reduce the de-anchoring indicator by moving the learning rate of the public closer to the optimal value.[22]

Third, to estimate the Fed's actual inflation target $\pi_t^\star$ and perceived inflation target π_t^P from US macro data, I generalize the precision based formulas for Bayesian estimation and model comparison of univariate correlated unobserved components models of Grant and Chan (2017) to the multivariate case. This generalization enables the use of precision based methods of Chan and Jeliazkov (2009) that are computationally more efficient than algorithms based on the Kalman filter and smoother. Moreover, the precision based methods directly deliver smoothed estimates that are best suited for the purpose of recovering historical relationships. Another advantage is that the Bayesian approach allows to explicitly

---

[21]When agents learn from inflation surprises and not from monetary policy, any change in their long-term inflation belief is considered undesirable and, thus, the degree of anchoring increases the smaller the learning rate is; see e.g. Jorgensen and Lansing (2022), Gáti (2022), Carvalho et al. (2022) and Lansing (2009).

[22]Other studies that estimate a constant learning rate are Erceg and Levin (2003) and Del Negro and Eusepi (2011). Carvalho et al. (2022) and Mertens and Nason (2020) estimate time-varying updating parameters for the inflation trend but their analysis omits an explicit role of the monetary policy and the inflation target, which are crucial for gauging credibility.

incorporate prior beliefs on structural parameters and the trajectories of the unobserved components. For example, it is straightforward to explicitly account for the implicit prior belief that $\pi_t^\star$ should be *close* to 2% from 2012 onwards.[23]

Results. The model is estimated on US macro data from 1962Q1 to 2018Q3. The best fitting model allows for five breaks in the variances and the learning rate. Interestingly, the variances of shocks to $\pi_t^\star$ and π_t^P show different patterns of variation across the regimes, a finding that would have been ruled out by the assumption of equal variances by Kozicki and Tinsley.

The estimates for $\pi_t^\star$ and π_t^P indicate that imperfect credibility is an important feature to fit the evolution of the Fed's monetary policy. Imperfect credibility was particularly important in the Volcker Disinflation due to the slow adjustment of agent's perception after the shift to a lower inflation target, and to a lesser extent in the aftermath of the 2008 Financial Crisis. A model with perfect credibility, i.e. $\pi_t^\star = \pi_t^p$ for all t, is strongly rejected by the data.

The estimated level of the Fed's inflation target $\pi_t^\star$ is close to 2% even before the 2012 announcement and it remains there since then. Re-estimating the model under an additional prior that restricts $\pi_t^\star$ to be 'close' to 2% does not deteriorate the model fit significantly. However, the de-anchoring indicator improves after the announcement. While the learning rate increases marginally, this increase is too small to contribute significantly to the improvement in anchoring. Instead, the improvement in anchoring is driven predominantly by a reduction of the variance of the transitory monetary policy shock, indicating that the Fed has followed its policy rule more closely since the announcement.

Moreover, the perceived target and observed long-term inflation expectations should not be equated directly, even though both capture similar economic concepts. A Bayesian model comparison shows that using long-term inflation expectations from the Survey of Professional Forecasters (SPF) as an imperfect measurement of the perceived target as in Chan et al. (2018) leads to a substantial deterioration in model fit.

Outline. The next section presents the learning mechanism of Kozicki and Tinsley, which serves as a starting point. It then proceeds to derive the optimal learning rate and

---

[23]Kozicki and Tinsley (2005) estimate the model with ML methods and report filtered estimates of $\pi_t^\star$ and π_t^P from the Kalman filter with no bands for inference. Following Kim and Kim (2022), Bayesian techniques should be preferred over maximum likelihood estimation for unobserved components models because they allow for overcoming the so-called 'pile-up' problem that can lead to a bias in the estimates of variances of the unobserved components.

implications for the evolution of the credibility gap and the degree of anchoring of the perceived target to the Fed's actual target. Section 3 embeds the learning mechanism into a multivariate correlated unobserved components model and derives analytical expressions for Bayesian estimation and model comparison in a precision based framework. Section 4 presents the results of the baseline estimation and alternative specifications to assess the importance of imperfect credibility. Section 5 summarizes and concludes.

2.2 An empirical macro model with asymmetric information

To assess the time-varying credibility of the Fed's inflation target I build on the asymmetric information model by Kozicki and Tinsley (2005). This section briefly revisits the learning mechanism of their model. The main components are the monetary policy rule and the learning rule that updates the perceived inflation target. The subsequent subsection derives results of the model for credibility and the optimal learning rate.

2.2.1 The learning mechanism of Kozicki and Tinsley (2005)

Kozicki and Tinsley assume that monetary policy sets the nominal interest rate i_t according to a Taylor-type feedback rule in response to deviations of the annual inflation rate $\bar{\pi}_t^4$ from the current value of the inflation target $\pi_t^\star$ and to a measure of the output gap $g_t - \bar{g}$. Moreover, it smooths past interest rate deviations i_{t-1} from the equilibrium level $\bar{i}_{t-1}$. The monetary policy rule reads

$$i_t = \bar{i}_t + \rho(i_{t-1} - \bar{i}_{t-1}) + \phi_{\pi,t}(\bar{\pi}_t^4 - \pi_t^\star) + \phi_y(g_t - \bar{g}) + \varepsilon_t^{MP}, \quad \varepsilon_t^{MP} \sim N(0, \sigma_{MP}^2)$$
$$(2.1)$$

with $\quad \bar{i}_t = rr + \pi_{t-1}^P.$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.2)$

The equilibrium level of the nominal rate $\bar{i}_t$ is determined by a long-term Fisher-type relation where rr denotes the natural real rate and π_{t-1}^P is the perceived inflation target from the previous period. As is common in the literature on time-varying inflation targets, $\pi_t^\star$ follows

an exogenous random walk, see e.g. Ireland (2007) and Cogley et al. (2010).[24]

$$\pi_t^\star = \pi_{t-1}^\star + \varepsilon_t^\star, \quad \varepsilon_t^\star \sim N(0, \sigma_\star^2) \tag{2.3}$$

While the monetary policy shock ε_t^{MP} only has a transitory effect in the policy interest rate, the target shock $\varepsilon_t^\star$ has a permanent effect.

Asymmetric information enters the model through the assumption that agents outside the central bank do not directly observe either the current value of the Fed's inflation target or the transitory shock. To form beliefs about the inflation target, which is the perceived inflation target π_t^P, agents have to solve a signal extraction problem. This involves forming expectations about the policy rate i_t^e and updating the perceived target π_t^P according to the surprise $i_t - i_t^e$. Since $\pi_t^\star$ is unknown, agents use their latest value of π_{t-1}^P to form interest rate expectations. Thus, agents' one period ahead expected level of the interest rate is given by

$$i_t^e = \bar{i}_t + \rho(i_{t-1} - \bar{i}_{t-1}) + \phi_\pi(\bar{\pi}_t^4 - \pi_{t-1}^P) + \phi_y(g_t - \bar{g})$$

After observing the actual level of the interest rate set via (2.1), agents' forecast error can be decomposed into contributions of transitory and permanent monetary policy components according to

$$i_t - i_t^e = \phi_\pi(\pi_{t-1}^P - \pi_t^\star) + \varepsilon_t^{MP}. \tag{2.4}$$

If monetary policy is more contractionary than expected, the decomposition in (2.4) implies that the corresponding forecast error $i_t - i_t^e > 0$ can be attributed to either a positive transitory monetary policy shock ε_t^{MP} or to agents' perceived target being higher than the actual inflation target, i.e. $\pi_{t-1}^P - \pi_t^\star > 0$. Since both ε_t^{MP} and $\pi_t^\star$ are unobservable, agents do not know the source of their forecast error. However, they can adjust their perceived target according to size and direction of the forecast error. Following Kozicki and Tinsley, I

---

[24]Here, I deviate from the original model in two aspects. First, I do not include a 'Volcker dummy' in the law of motion for $\pi_t^\star$ to account for the large monetary policy shift that occurred between 1979 and 1983. Instead, changes in monetary policy are captured by allowing for changes in the variances of the shocks. For the Volcker Disinflation, this has the advantage that, while also allowing for larger target changes, this special regime can also be reflected in a larger variance of temporary monetary policy shocks, due to non-borrowed reserves targeting. Second, Kozicki and Tinsley allow the inflation target to react to temporary cost-push shocks, which seems implausible for most of the sample period. Moreover, since this is not at the heart of the current analysis, I abstract from this feature and assume that the inflation target is a purely exogenous random walk as in (2.3).

assume that agents employ the following learning rule to update the perceived target

$$\pi_t^P = \pi_{t-1}^P - \delta(i_t - i_t^e) + \varepsilon_t^P, \quad \varepsilon_t^P \sim N(0, \sigma_P^2). \tag{2.5}$$

The learning rate δ governs the relative weight that agents attach to new information in the expectation formation process. It can also be interpreted as the amount of attention that agents devote to monetary policy actions. If monetary policy is more contractionary than expected, agents revise π_t^P downward. The perceptions shock ε_t^P is an exogenous source of variation in π_t^P and allows for deviations from the learning rule. I deviate from Kozicki and Tinsley (2005) in that I do not impose the restriction $\sigma_P^2 = \sigma_\star^2$ that was originally made for technical reasons but is hard to justify on economic grounds. The Bayesian estimation allows for replacing this restriction with a reasonable prior that still allows the two variances to be different.[25]

Substituting (2.4) into (2.5) yields the law of motion for π_t^P as

$$\pi_t^P = (1 - \delta\phi_\pi)\pi_{t-1}^P + \delta\phi_\pi\pi_t^\star - \delta\varepsilon_t^{MP} + \varepsilon_t^P. \tag{2.6}$$

If $\phi_\pi > 0$ and $\delta > 0$, (2.6) implies that π_t^P is cointegrated one for one with the inflation target $\pi_t^\star$ with adjustment coefficient $\phi_\pi\delta$ and perceptions will eventually converge to the Fed's target in the absence of shocks. For the observed data, this implies that $\pi_t^\star$ is the common trend of inflation π_t and the monetary policy interest rate i_t.

2.2.2 The credibility gap

The asymmetric information structure of the model prevents a first-best outcome of a perfectly credible inflation target $\pi_t^\star = \pi_t^P$ with zero variance for all t. To overcome asymmetric information in practice, central banks use their communication with the public as an additional tool for achieving credibility for their inflation targets. To that end, the communication strategies are often subject to change, while the general monetary policy framework remains largely unchanged. However, a change in the communication strategy should affect the weight of monetary policy in the public perceptions formation process, i.e. the learning rate δ. In fact, the learning rate δ can achieve a second-best outcome that is associated with minimum volatility and persistence of deviations of π_t^P from $\pi_t^\star$. This section first defines the credibility gap $\pi_t^P - \pi_t^\star$ and then derives a value for the learning rate

---

[25]As I am additionally allowing for different volatility regimes in the estimation, relaxing this restriction is even more important because it would *a priori* imply the same variance changes in target- and perceived target shocks in each regime.

$\delta^{\star}_{t,s}$ that is conditionally optimal from a central bank point of view that seeks to minimize expected credibility gaps for horizons $s = 1$ and $s = \infty$. These two cases allow to illustrate the intuition of optimal learning and serve as benchmarks for interpreting the empirically estimated learning in section 2.4.

The model allows a law of motion for credibility, defined as „the difference between the policymaker's plan and the public's beliefs about those plans" Cukierman and Meltzer (1986, p.1106). This definition corresponds to the deviation of the perceived target from the Fed's actual target $\pi_t^P - \pi_t^{\star}$, that is bad credibility drives a wedge between $\pi_t^{\star}$ and π_t^P. To stress that larger magnitudes of $|\pi_t^P - \pi_t^{\star}|$ imply weaker credibility, it is referred to as the *credibility gap*. Subtracting $\pi_t^{\star}$ from (2.6) and inserting (2.3) yields the law of motion for the credibility gap:[26]

$$\pi_t^P - \pi_t^{\star} = (1 - \delta\phi_{\pi})(\pi_{t-1}^P - \pi_{t-1}^{\star}) + \varepsilon_t^P - \delta\varepsilon_t^{MP} - (1 - \delta\phi_{\pi})\varepsilon_t^{\star} \qquad (2.7)$$

Again, if $0 < \phi_{\pi}\delta < 1$, the credibility gap converges to zero in absence of shocks implying that π_t^P is anchored at $\pi_t^{\star}$ *in the long run* in the sense of the long-run anchoring criterion of Nautz et al. (2019). For the remainder of this section, assume that this condition is satisfied. The estimation results show that the upper bound is empirically not binding. Given that π_t^P is anchored in the long run, the volatility and persistence of the credibility gap are two aspects a central bank might be concerned about.[27]

2.2.3 Optimal learning

The volatility and persistence of the credibility gap are naturally summarized by expected squared deviations of π_{t+s}^P from $\pi_{t+s}^{\star}$ at a medium- to long-term horizon s. To emphasize the role of these aspects for the degree of anchoring of π_t^P to $\pi_t^{\star}$, those deviations are referred to as the *de-anchoring indicator $DAI_{t,s}$*, see Definition 1.

Definition 1 (The De-Anchoring Indicator) *Conditional on the information set as of period t, denoted $\mathcal{I}_t$, the degree of de-anchoring of π_{t+s}^P from $\pi_{t+s}^{\star}$ s periods into the future is measured by*

$$DAI_{t,s} := \mathrm{E}\left[\left(\pi_{t+s}^P - \pi_{t+s}^{\star}\right)^2 \Big| \mathcal{I}_t\right]$$

---

[26]For the VECM describing the evolution $\pi_t^{\star}$ and π_t^P, this reformulation corresponds to the more general transformation in Carvalho and Harvey (2005, p.278) that uses eigenvectors and eigenvalues.

[27]This is akin to the criteria considered for the anchoring of long-term inflation expectations, see e.g. Doh and Oksol (2018).

Given a set of monetary policy coefficients, the rate of learning δ plays a central role for the $DAI_{t,s}$. Therefore, a central bank that seeks to maintain credibility in the medium to long term prefers a value $\delta_{t,s}^{\star}$ that minimizes $DAI_{t,s}$:

$$\delta_{t,s}^{\star} = \operatorname*{argmin}_{\delta \in \left(0, \frac{1}{\phi_\pi}\right)} DAI_{t,s}. \tag{2.8}$$

Since $DAI_{t,s}$ is the long-run variance of the credibiltiy gap when $s \to \infty$, for this case the problem resembles the criterion for optimal learning proposed by Lansing (2009). The value $\delta_{t,s}^{\star}$ is optimal for a central bank that chooses the Taylor rule coefficient ϕ_π based on considerations about the trade-off between inflation and output stabilization and maximizes anchoring subject to that policy rule. The announcement of the 2% inflation target by the Federal Reserve can be understood as such an attempt to maximize anchoring of long-term inflation expectations *conditional* on the monetary policy regime already in place. If such communication tools have the desired anchoring-effect, they should move δ closer to $\delta_{t,s}^{\star}$.

Using the law of motion for the credibility gap allows to derive an analytical expression for the $DAI_{t,s}$, see Proposition 1. The proof is contained in Appendix 2.A.

Proposition 1 (Model implied $DAI_{t,s}$) *The credibility gap in equation* (2.7) *implies*

$$DAI_{t,s} = \left(\pi_t^P - \pi_t^{\star}\right)^2 (1 - \delta\phi_\pi)^{2s} + \left(1 - (1 - \delta\phi_\pi)^{2s}\right) \frac{\left(\delta^2\sigma_{MP}^2 + \sigma_*^2 (1 - \delta\phi_\pi)^2 + \sigma_P^2\right)}{1 - (1 - \delta\phi_\pi)^2} \tag{2.9}$$

For the case $s = 0$, where the central bank does not care about future deviations of the perceived target from $\pi_t^{\star}$, $DAI_{t,s}$ collapses to the squared current credibility gap. By contrast, if the central bank cares also about future deviations, the degree of de-anchoring is not reflected adequately by only the current credibility gap. As central banks are forward looking, $s = 0$ is hardly an empirically relevant choice. Therefore, this section proceeds to analyze the degree of anchoring for the more plausible case of $s > 0$.

In principle, the horizon s can be chosen to match the central bank's definition of the 'medium term'. However, the mechanics of the $DAI_{t,s}$ are best illustrated by two extreme cases $s \to \infty$ and $s = 1$ which allow to simplify $DAI_{t,s}$ considerably, see Corollary 1.

Corollary 1 (Special cases for $DAI_{t,s}$) *For $s \to \infty$ and $s = 1$, $DAI_{t,s}$ simplifies to*

$$DAI_\infty = \frac{\delta^2\sigma_{MP}^2 + (1 - \delta\phi_\pi)^2\sigma_\star^2 + \sigma_P^2}{1 - (1 - \delta\phi_\pi)^2} \tag{2.10}$$

and

$$DAI_{t,1} = \left(\pi_t^P - \pi_t^\star\right)^2 (1 - \delta\phi_\pi)^2 + \delta^2\sigma_{MP}^2 + \sigma_*^2 (1 - \delta\phi_\pi)^2 + \sigma_P^2. \qquad (2.11)$$

Inserting the expression for DAI_∞ from (2.10) into the expression for $DAI_{t,s}$ in (2.9) yields

$$DAI_{t,s} = \left(\pi_t^P - \pi_t^\star\right)^2 (1 - \delta\phi_\pi)^{2s} + \left(1 - (1 - \delta\phi_\pi)^{2s}\right) DAI_\infty. \qquad (2.12)$$

From this, it is easy to see that $DAI_{t,s}$ converges to DAI_∞ monotonically as $s \to \infty$. Moreover, $DAI_{t,s}$ approaches DAI_∞ from above (below) if the current squared credibility gap is larger (smaller) then DAI_∞. Given long run anchoring, both DAI_∞ and $DAI_{t,1}$ are decreasing in the Taylor rule coefficient on inflation ϕ_π due to the decline in persistence of the credibility gap. This intuitive finding is also in line with the theory of anchored inflation expectations of Gáti (2022), who concludes that a more aggressive central bank reaction to inflation deviations anchors long-run inflation beliefs. Moreover, both de-anchoring indicators are increasing in the variances σ_P^2, σ_{MP}^2 and, $\sigma_\star^2$.

Corollary 2 summarizes the effect of a change in δ on DAI_∞ and $DAI_{t,1}$. Due to the nature of asymmetric information, a higher learning rate does not necessarily improve anchoring. This is because an increase in δ has two opposite effects. On the one hand, it improves anchoring by reducing the persistence of the credibility gap, reflected in the right hand side of the conditions in Corollary 2. On the other hand, a higher δ also increases the impact of transitory monetary policy shocks on the credibility gap, represented by the left hand side of the conditions. Thus, a higher δ only improves anchoring if the reduction in persistence dominates the higher impact of the monetary policy shock on the credibility gap. Moreover, the strength of these opposite effects differs between the horizons s. For example, for $s \to \infty$, the reduction in persistence is more likely to dominate if σ_t^P is large. For $s = 1$, this is the case if the current credibility gap is large.

Corollary 2 *Given $0 < \phi_\pi$ and $0 < \delta \leq \frac{1}{\phi_\pi}$, for $s \to \infty$ we have*

$$\frac{\partial \, \text{DAI}_\infty}{\partial \delta} \begin{cases} \leq 0, & for \quad \delta^2\sigma_{MP}^2 \leq (\sigma_*^2 + \sigma_P^2)(1 - \delta\phi_\pi) \\ > 0, & for \quad \delta^2\sigma_{MP}^2 > (\sigma_*^2 + \sigma_P^2)(1 - \delta\phi_\pi) \end{cases}$$

and for $s = 1$

$$\frac{\partial \text{DAI}_{t,1}}{\partial \delta} \begin{cases} \leq 0, & for \quad \delta^2 \sigma_{MP}^2 \leq \left(\sigma_*^2 + \left(\pi_t^P - \pi_t^\star \right)^2 \right) \phi_\pi \left(1 - \delta \phi_\pi \right) \\ > 0, & for \quad \delta^2 \sigma_{MP}^2 > \left(\sigma_*^2 + \left(\pi_t^P - \pi_t^\star \right)^2 \right) \phi_\pi \left(1 - \delta \phi_\pi \right) \end{cases}$$

Proposition 2 (The optimal learning rates) *For $s \to \infty$ and $s = 1$,*

$$\delta_\infty^\star = \frac{\sqrt{(\sigma_*^2 + \sigma_P^2)(\phi_\pi^2 \sigma_*^2 + \phi_\pi^2 \sigma_P^2 + 4\sigma_{MP}^2)} - \phi_\pi (\sigma_*^2 + \sigma_P^2)}{2\sigma_{MP}^2} \tag{2.13}$$

and

$$\delta_{t,1}^\star = \frac{\phi_\pi \left(\left(\pi_t^P - \pi_t^\star \right)^2 + \sigma_*^2 \right)}{\phi_\pi^2 \left(\left(\pi_t^P - \pi_t^\star \right)^2 + \sigma_*^2 \right) + \sigma_{MP}^2}. \tag{2.14}$$

solve the problem in (2.8).

Solving (2.8) yields the optimal learning rates, see Proposition 2. The proof is contained in Appendix 2.A. The optimal learning rates $\delta_\infty^\star$ and $\delta_{t,1}^\star$ depend differently on the shock variances and the Taylor rule parameter. Thus, to judge whether a higher or lower value of δ is needed to improve anchoring at a specific horizon, it is useful to compare any estimated value for δ with both extreme cases $\delta_\infty^\star$ and $\delta_{t,1}^\star$. For example, in the case of a perfectly credible inflation target at time t, i.e. $\pi_t^P - \pi_t^\star = 0$, and a central bank that follows the Taylor rule almost exactly, i.e. $\sigma_{MP}^2 \to 0$, the optimal learning rate $\delta_{t,1}^\star$ approaches its upper bound $\frac{1}{\phi_\pi}$ under long-run anchoring. The maximum learning speed is optimal in this case because every forecast error agents make in forecasting the interest rate i_t originates from a change in $\pi_t^\star$. If the central bank would not follow the Taylor rule in setting its interest rate policy, i.e. $\sigma_{MP}^2 \to \infty$, the interest rate contains no information about the inflation target and, consequently, a learning rate of close to zero would be optimal. However, in the empirically relevant case, where the central bank follows the Taylor rule approximately, a modestly positive learning rate is optimal.

In contrast, models that assume long-term inflation beliefs are formed based on inflation surprises exclusively imply that a learning rate of close to zero maximizes anchoring because every movement in long-term inflation beliefs is undesirable by definition; see for example Carvalho et al. (2022). The results of this section show that, when agents learn from monetary policy instead of inflation surprises, a non-zero learning rate can be optimal.

2.2.4 US Monetary policy regimes and the learning rate

To estimate a meaningful learning parameter δ, it is necessary to capture monetary policy adequately in the estimation. The SVAR literature on US monetary policy largely agrees that the variances of the structural shocks have changed across the different monetary regimes, while the coefficients of the policy reaction function remained remarkably stable in post war data (Sims and Zha 2006b; Belongia and Ireland 2016). Changes in volatilities are of particular interest for the anchoring of the perceived target to the actual target because the optimal learning rate and the de-anchoring indicators depend on the volatilities. To account for potential changes in volatilities, I follow Brunnermeier et al. (2021) and allow the structural shock variances to change between break dates suggested by the literature.

In addition to the break dates of Brunnermeier et al. (2021), I allow for a break in January 2012 for the announcement of the official 2% inflation target. Intuitively, the variances of changes in the inflation target $\sigma_\star^2$ and deviations from the Taylor rule σ_{MP}^2 may have decreased after the announcement.

Even under constant variances the optimal $\delta_{t,1}^\star$ is time-varying because it depends on the current value of the credibility gap. In contrast, $\delta_\infty^\star$ is constant because it "sees through" current credibility gaps that are only temporary under long-run anchoring. However, time-variation in the variances of the shocks can imply time-variation in $\delta_\infty^\star$. Since the optimal learning rates depends on the shock variances, I also allow the learning rate δ to change at the break dates.[28]

Studies focusing on learning from inflation surprises document time-variation in the estimated learning rates, see e.g. Carvalho et al. (2022), Jorgensen and Lansing (2022) and Gáti (2022). In contrast, the learning mechanism of Kozicki and Tinsley assumes that agents learn from monetary policy surprises. To the best of my knowledge, the present study is the first to estimate the time-varying learning gain in a monetary policy-based learning mechanism, and hence, adds an important perspective the literature. For example, if the 2012 announcement had the desired effect on credibility and anchoring, it should have led to a decline in the variance of the target- and transitory monetary policy shocks and shifted the learning rate δ closer to its optimal value.

---

[28]The breaks in the signal-to-noise ratio due to breaks in variances imply changes in the optimal learning gain not only from the anchoring analysis, but also from a Kalman filter, i.e. optimal forecasting, perspective.

2.3 Bayesian estimation

To estimate the inflation target $\pi_t^\star$ and the perceived inflation target π_t^P, the structural equations of the previous section are mapped to state space form, This form yields a multivariate unobserved components model with correlated errors. The vector of $n = 3$ macroeconomic variables $y_t = [g_t, \pi_t, i_t]'$ consists of a measure of the output gap g_t, inflation π_t and the central bank interest rate i_t. The data vector y_t is decomposed into the $r = 2$ common trends, collected in the vector $\tau_t = [\pi_t^P, \pi_t^\star]'$. The deviations of the variables from the trends are denoted $c_t = [\hat{g}_t, \hat{\pi}_t, \hat{i}_t]'$ and are assumed to be stationary. The full observation equation linking the observables to the state vectors is

$$y_t = \tilde{\gamma}_y + \Gamma_0 \tau_t + \Gamma_1 \tau_{t-1} + c_t \tag{2.15}$$

with

$$\tilde{\gamma}_y = \begin{bmatrix} \bar{g} \\ 0 \\ rr \end{bmatrix} \qquad \Gamma_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \Gamma_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

where $\bar{g}$ is the constant average of the output gap that might be different from zero in a particular sample period and rr is the real interest rate. Note that only π_t^P enters directly into the inflation equation; see second row of Γ_0. Moreover, π_t^P appears with a lag in the interest rate equation; see third row of Γ_1. The state equations governing the evolution of the cycles c_t and the trends τ_t are

$$Ac_t = B_1 c_{t-1} + ... + B_p c_{t-p} + \lambda_0 \tau_t + ... + \lambda_q \tau_{t-q} + e_t, \qquad e_t \sim N(0, \Sigma_t) \tag{2.16}$$

$$\tau_t = F_t \tau_{t-1} + J_t e_t + O_t u_t, \qquad\qquad u_t \sim N(0, \Omega_t) \tag{2.17}$$

where $u_t = [\varepsilon_t^P, \varepsilon_t^\star]'$ and $e_t = [\varepsilon_t^g, \varepsilon_t^\pi, \varepsilon_t^{MP}]'$. The t subscript of the diagonal variance matrices Σ_t and Ω_t indicates the dependence on the US monetary policy regimes. The coefficient matrices F_t, J_t, and O_t also have a t subscript because they depend on the learning rate δ, which is also allowed to vary across the regimes. For brevity, the exact definitions the model matrices are relegated to Appendix 2.B. Suffice to mention that A is lower triangular with unit diagonal. The coefficients in the nth row of A, the B_js and the λ_js obey linear restrictions such that the interest rate equation equals the Taylor rule in (2.1). After substituting the decomposition $\pi_t = \pi_t^P + \hat{\pi}_t$ from (2.15) into the Taylor rule, the λ_j matrices account for how $\pi_t^\star$ and π_{t-i}^P for $i = 0, ..., 3$ affect interest rate deviations $\hat{i}_t$. Since

$\pi_t^\star$ and π_{t-i}^P do not enter any other equation for c_t, the λ_j contains only zeros elsewhere. Equation (2.17) stacks laws of motion for $\pi_t^\star$ and π_t^P in (2.3) and (2.6). Since the cycle shocks e_t show up in the equations for τ_t and c_t, I refer to the model defined by equations (2.15)-(2.17) as multivariate correlated unobserved components (MCUC) model.

2.3.1 Priors

I use natural conjugate priors where possible to allow for an efficient estimation. This means normal distributions as priors for all slope coefficients and initial values of the unobserved components, and inverse gamma distributions for all variances. For the shock variances of the stationary cycles σ_g^2, σ_π^2, and σ_{MP}^2, I use inverse gamma distributed priors with a mean of 0.5 and 6 degrees of freedom. For the shock variances of the trends $\sigma_\star^2$ and σ_P^2, I use inverse gamma distributed priors with a smaller mean of 0.05 and 5 degrees of freedom. The smaller mean is justified because changes in the inflation target or the perceived target can be expected to be smaller on average than business cycle shocks. Still, the prior is not overly restrictive. For example, at the prior mean, 95% of the changes in $\pi_t^\star$ are smaller than 0.44 in absolute value. For the Taylor rule coefficients ρ, ϕ_π, and ϕ_g, I use informative normal priors with mean 0.7, 0.45, and 0.15 and variances of 0.05 each that are reminiscent of priors for these parameters from Smets and Wouters (2007). The use of informative priors is justified because these are structural parameters with a clear economic meaning. The other slope coefficients in the B_i matrices of all other cyclical equations have unspecific and wide normal priors with zero mean and unit variance. For the learning rate δ, I use a beta prior of the form $\delta \sim beta(a_\delta, b_\delta)$ with $a_\delta = 4$ and $b_\delta = 16$ implying a mean of 0.2. This prior restricts $0 < \delta < 1$ and ensures that the long-run anchoring criterion is obeyed. While the lower bound also ensures the correct sign for δ, the upper bound is not empirically relevant.

2.3.2 Posterior simulation and marginal likelihood computation

The model is estimated using Bayesian methods. The use of Bayesian methods has the advantage that prior beliefs on structural parameters and the trajectories of unobserved components can be explicitly taken into account. Moreover, the different priors on the variances of the cycle and trend shocks add to the identification of the unobserved components without imposing hard restrictions.[29]

---

[29]Kozicki and Tinsley (2005) estimate the model with ML methods and report filtered estimates of $\pi_t^\star$ and π_t^P from the Kalman filter with no bands for inference. Following Kim and Kim (2022), Bayesian

The posterior of the model is simulated with a Gibbs sampler. The Gibbs sampler approximates the posterior by iteratively generating draws from the conditional posterior distribution of the unknown parameters and the states τ_t, reminiscent of an Expectation Maximization algorithm. For the purpose of recovering in-sample relations, smoothed estimates of τ_t are most appropriate because they use all the available sample information. The most efficient way to draw τ_t from its smoothing distribution is by use of the precision sampler of Chan and Jeliazkov (2009). Unfortunately, the formulas for the precision sampler for correlated unobserved components models are only available for the univariate case, see Grant and Chan (2017). Therefore, to estimate the MCUC model, I generalize the formulas to the multivariate case. The individual steps of the algorithm are detailed in Appendix 2.D. To ensure convergence of the Gibbs sampler I discard the first 5000 draws as burn-in sample. All results are based on the 20000 draws following the burn-in.

To efficiently compute the marginal likelihood, required for Bayesian model comparison, I also generalize the analytical computation of the integrated likelihood of Grant and Chan (2017) to the multivariate case. The marginal likelihood is obtained by numerically integrating out the unknown coefficients from the integrated likelihood. This requires many evaluations of the integrated likelihood, which is greatly facilitated by a closed form expression that can be evaluated quickly. Using this analytical expression, I obtain the marginal likelihood via the cross-entropy method of Chan and Eisenstat (2015). The cross-entropy method is an importance sampling procedure that requires specifying distribution families for the proposal densities for all parameters. Following Chan and Eisenstat (2015), I use proposal densities from the same families as the prior densities for each coefficient. I use 10 batches with 10000 draws each to compute the marginal likelihood. Since the cross entropy method yields a numerical approximation, there is also a small error. To gauge the approximation error, I compute a numerical standard error (NSE) for the marginal likelihood estimates across the 10 batches.

For brevity, the formulas and derivations are contained in 2.C. The formula for the integrated likelihood of the MCUC can also be used in maximum likelihood estimation.

---

techniques should be preferred over maximum likelihood estimation for unobserved components models because they allow for overcoming the so-called 'pile-up' problem that can lead to a bias in the estimates of variances of the unobserved components.

2.4 Credibility of the Fed's inflation target from 1962 to 2018

I estimate the model on three quarterly US time series from 1962Q1 to 2018Q3: the output gap g_t obtained from the Congressional Budget Office, the annualized quarterly inflation rate of the Personal Consumption Expenditure π_t, and the Fed Funds rate i_t. From 2009 to 2015, while the zero lower bound was binding, i_t is equal to the shadow rate of Wu and Xia (2016), which captures the unconventional monetary measures taken during that time.[30] Before interpreting the results, I determine the monetary policy break dates out of the candidate break dates that yield the best fit.

2.4.1 Determining the monetary policy break dates

To fit the US data and find the most relevant break dates, I compare the marginal likelihoods from models estimated with different sets of break dates from Brunnermeier et al. (2021) plus a break in January 2012 for the inflation target announcement. Table 2.4.1 confirms that variance changes are a relevant feature to fit the data. The best fitting model with the largest marginal likelihood has five break dates in total; see row 6. The regime breaks refer to the first month of the new regime. The five breaks that yield the best fit mark the beginnings of the the Stagflation period in January 1973, the regime change in October 1979 shortly after Paul Volcker was appointed chairman of the Fed, the end of the monetary targeting in January 1983, the onset of the financial crisis in January 2008, and the announcement of the inflation target in January 2012.

Allowing for these breaks in volatilities and the learning rate, the constant Taylor rule coefficients also have the expected sign; see Table 2.4.2. Moreover, the implied long-run response to inflation deviations from target exceeds unity and obeys the Taylor principle. The smoothing coefficient ρ is relatively low. Since time-variation in $\pi_t^\star$ and imperfect credibility, two elements that are absent in conventional specifications, also capture persistence in the interest rate, this Taylor rule requires only a smaller smoothing parameter.

The typical response to a temporary monetary policy ε_t^{MP} shock implied by the model is also plausible, see Figure 2.4.1. A 25 basis points hike in the fed funds rate leads to a decline in the output gap and a delayed decrease in inflation after a small price puzzle. The perceived target decreases modestly after the shock and, by construction, the inflation

---

[30]All series are obtained from economic database of the St. Louis Fed, FRED. Only the shadow rate, which is obtained from Cyntia Wu's homepage at `https://sites.google.com/view/jingcynthiawu/shadow-rates`.

Table 2.4.1: Log marginal likelihoods for various sets of regime break dates

Start date of new regime						
Jan-1973	Oct-1979	Jan-1983	Jan-1990	Jan-2008	Jan-2012	$\log ML$ (NSE)
						-965.6 (0.67)
					✓	-960.29 (0.42)
	✓	✓			✓	-927.14 (1.5)
✓	✓	✓			✓	-922.64 (1.11)
✓	✓	✓	✓		✓	-926.32 (1.29)
✓	✓	✓		✓	✓	-907.93 (1.22)
✓	✓	✓	✓	✓	✓	-913.73 (1.27)

Notes: Log marginal likelihoods are computed using the cross-entropy method of (Chan and Eisenstat 2015) with 10 runs of 10000 importance sampling draws each. Numerical standard error (NSE) across the 10 runs in parenthesis. Strength of evidence for differences in log ML according to Kass and Raftery (1995): $0 < \Delta \log ML < 1$: not worth mentioning, $1 < \Delta \log ML < 3$: positive, $3 < \Delta \log ML < 5$: strong, $5 < \Delta \log ML$: very strong. Additional break dates of Belongia and Ireland (2016) in Jan-2000 and Jan-1984 instead of Jan-1983 did not increase the fit log ML.

target remains constant.[31]

Since the Taylor rule and the impact of the monetary policy shock of the best fitting model are plausible, the results in the following sections are based on the same set of regime breaks. The next section considers the potential time-varying credibility.

2.4.2 The time-varying credibility of $\pi_t^\star$

Figure 2.4.2 shows the estimated paths for π_t^P and $\pi_t^\star$ and the credibility gap. It is obvious that π_t^P and $\pi_t^\star$ do not coincide over large parts of the sample and, thus, that credibility has not always been perfect. In fact, imperfect credibility was a problem during the Volcker

---

[31]Note the inflation decreases on impact because the perceived target π^P decreases on impact.

Table 2.4.2: Taylor rule coefficients of the best fitting model.

	ϕ_g	ϕ_π	ρ
mean (5% 95%)	0.25 (0.14 0.36)	0.51 (0.33 0.68)	0.63 (0.5 0.76)

Notes: The posterior means of the implied long-run responses to inflation and the output gap are $\frac{\phi_\pi}{1-\rho} = 1.11$ and $\frac{\phi_g}{1-\rho} = 0.70$.

Figure 2.4.1: Impulse responses to a monetary policy shock
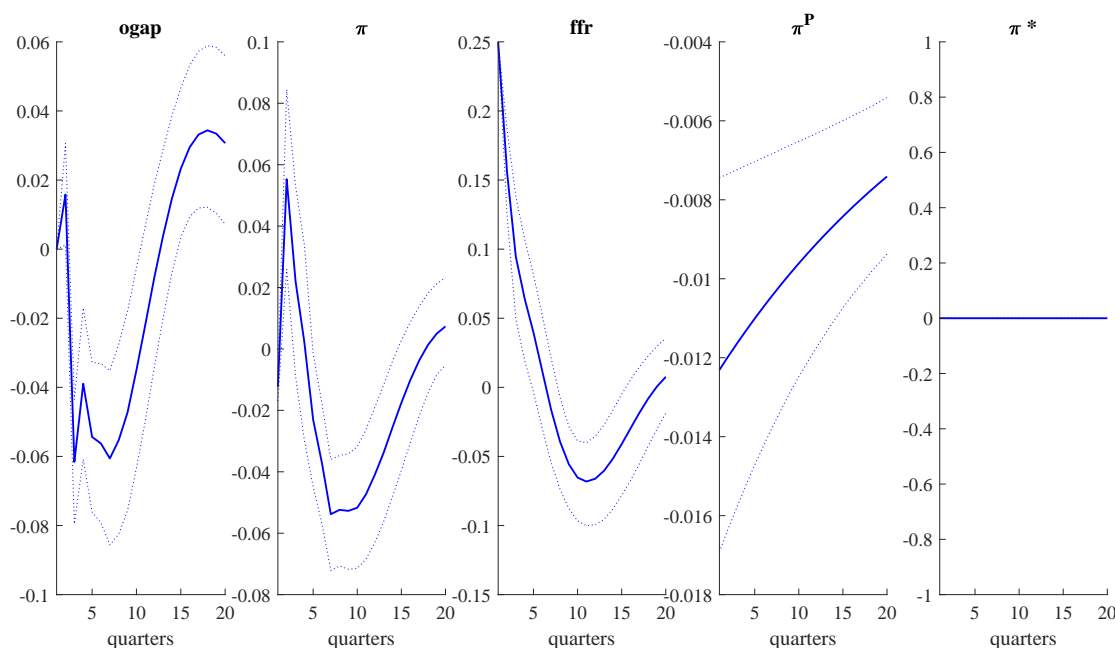
Notes: Posterior means (solid) 16% and 84% posterior quantiles (dashed lines). The shock is normalized to increase the fed funds rate by 25 basis points. To show the typical response implied by the model, the impulse responses are averaged over the values of the learning rate δ from the different monetary policy regimes.

Disinflation, and the Great Moderation, and, to a lesser extent, in the aftermath of the 2008 Financial Crisis. Until the mid-1970s, both, π_t^P and $\pi_t^\star$ increase steadily. Moreover, this rise is largely simultaneous indicated by the insignificant credibility gap during that period. Through the lens of the Taylor rule, a rise of $\pi_t^\star$ reflects the fact the Fed has not rigorously enforced low inflation with the 'go-stop'-type policy; see e.g. Goodfriend (2004). In the late 1970s and early 1980s, the so-called Volcker Disinflation, an apparent regime change takes place: $\pi_t^\star$ drops sharply to below 2% reflecting the Fed taking on the fight against high inflation. In contrast to the preceding simultaneous rise of π_t^P and $\pi_t^\star$, the drop in $\pi_t^\star$ clearly leads the decline in inflation and the gradual decline of π_t^P through the Great Moderation period. This results in a large positive credibility gap. The 68% probability

Figure 2.4.2: The evolution of π_t^P, $\pi_t^\star$ and the credibility gap
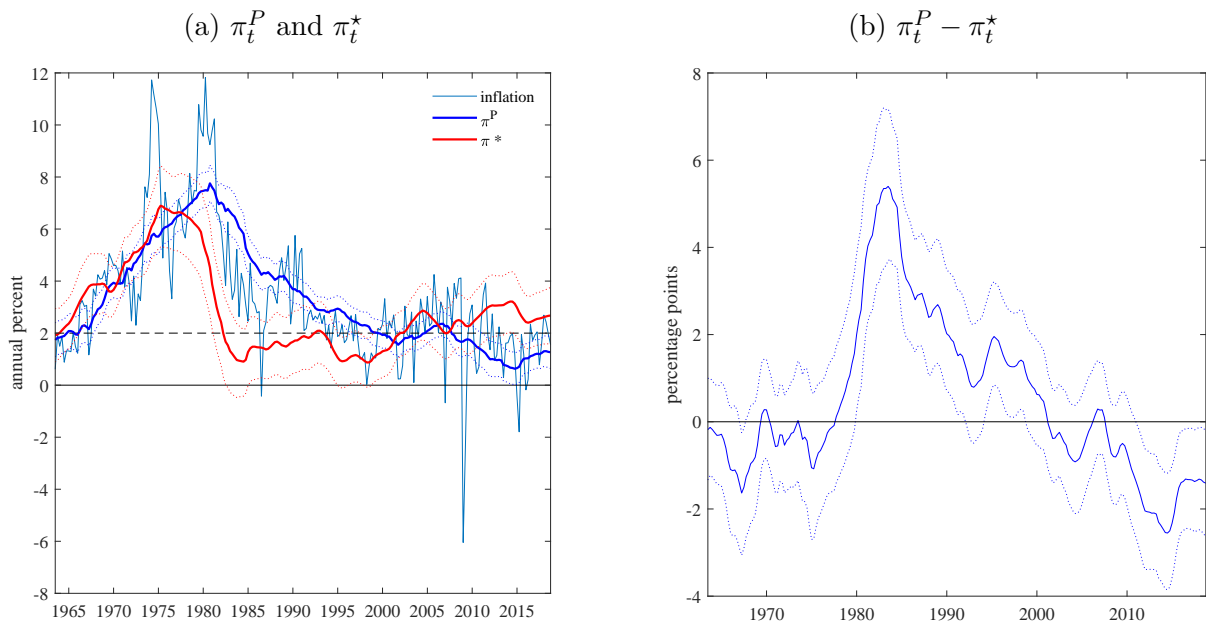
(a) π_t^P and $\pi_t^\star$ (b) $\pi_t^P - \pi_t^\star$

Notes: Horizontal lines indicate zero (solid) and 2% (dashed). Dotted lines are 16% and 84% posterior quantiles.

bands for the credibility bands contain zero again in the late 1990s.

Note that since the drop of $\pi_t^\star$ in the early 1980s, 2% is always contained in the 68% credibility bands.[32] This suggests that the 2012 announcement was a mere change in communication rather than a shift in the actual inflation target. However, in the aftermath of the Financial Crisis, the posterior mean of $\pi_t^\star$ increases to above 2%. This rise can be attributed partly to the unconventional monetary policy measures taken, which are reflected in the shadow rate. It may, despite the announced target, indicate a higher tolerance for inflation vis-a-vis other goals of the Fed. At the same time π_t^P declines below 2%. As a result, zero remains outside the probability band of the credibility gap through the sample end in 2018Q3. However, as shown in the previous sections, a central bank that cares about future credibility should also take into account the regime dependent volatilities and the learning rate becasue they determine the degree of de-anchoring at medium- to long horizons. Therefore, the next section considers those aspects jointly.

[32]However, $\pi_t^\star$ from Kozicki and Tinsley (2005) drops into negative territory at the end of the Volcker Disinflation just to increase sharply thereafter. While this excessively sharp drop can be due to their 'Volcker dummy' in the law of motion for $\pi_t^\star$, their estimate is much more volatile than the one in Figure 2.4.2, also in periods other than 1979Q4 to 1982Q4. This is reflected in their relatively large estimate for $\sigma_\star^2 = 0.23$ wich is more than twice the size of the maximum of 0.11 that I obtain outside the period from 1973Q1 to 1979Q3; see Table 2.4.3.

2.4.3 Learning and de-anchoring in different monetary regimes

The left panel of Figure 2.4.3 plots the estimated and optimal learning rates $\delta_{t,1}^\star$ and $\delta_\infty^\star$ along with the estimated δ across the regimes. Several stark observations emerge:

First, the estimated δ is lower than both optimal learning rates, in the entire sample period, indicating that agents update their beliefs too slowly.

Second, the profile of δ follows the profile of the optimal rates $\delta_\infty^\star$ and $\delta_{t,1}^\star$ across the regimes. For example, all three drop under the Financial Crisis regime. The drop of $\delta_\infty^\star$ is mostly driven by the increase of σ_{MP}^2 from 2008Q1 to 2011Q4 while the other variances that determine $\delta_\infty^\star$, i.e. σ_P^2 and $\sigma_\star^2$, remain almost unchanged; see Table 2.4.3. The increase in σ_{MP}^2 most likely reflects the modest decline in the interest rate compared to the sharp temporary drop in inflation. The fact that δ also drops along with $\delta_\infty^\star$ in this period suggests that agents were well aware of this.

Third, the learning rate was higher before the Great Recession than after. This finding is generally in line with findings of Carvalho et al. (2022), who report a lower learning rate in the more recent period. Also, Jorgensen and Lansing (2022) find that the learning rate drops even further after the Great Recession. However, the implications are different: In their models, agents learn from *inflation surprises* and, thus, a lower learning rate is associated with better anchoring of inflation beliefs because shocks to inflation will not lead big movements in long-term beliefs. In contrast, when agents learn from monetary policy, a learning rate of close to zero is not optimal because the link between the actual target and the perceived target becomes weaker. This is reflected in both $DAI_{t,1}$ and DAI_∞, in the middle and right panel, peaking in the aftermath of the 2008 Financial Crisis. Hence, considering only the effect of inflation surprises on agent's long-term inflation beliefs might overstate the degree of anchoring.

To better understand the dynamics of the $DAI_{t,1}$ and DAI_∞ through the entire sample, it is helpful to also take into account the changing volatilities in Table 2.4.3. In line with the narrative account of the Great Inflation, the estimated de-anchoring indicators both peak in the early 1980s and then decline. These peaks are driven by the high volatilities of temporary monetary policy, and inflation target shocks, σ_{MP}^2 and $\sigma_\star^2$.

Both de-anchoring indicators peak a second time in the 2008 recession. For $DAI_{t,\infty}$, the peak in the Volcker Disinflation is lower than the peak in the 2008 Recession, whereas the opposite is true for $DAI_{t,1}$. Taken at face value, $DAI_{t,\infty}$ would suggest that the de-anchoring was as severe as never before in 2008. However, this interpretation might be misleading because the $DAI_{t,\infty}$ completely ignores the current credibility gap $\pi_t^P - \pi_t^\star$, which is much lower after the Financial Crisis than during the Volcker Disinflation. Therefore, a high

Figure 2.4.3: Estimated and optimal learning rate, and de-anchoring indicators
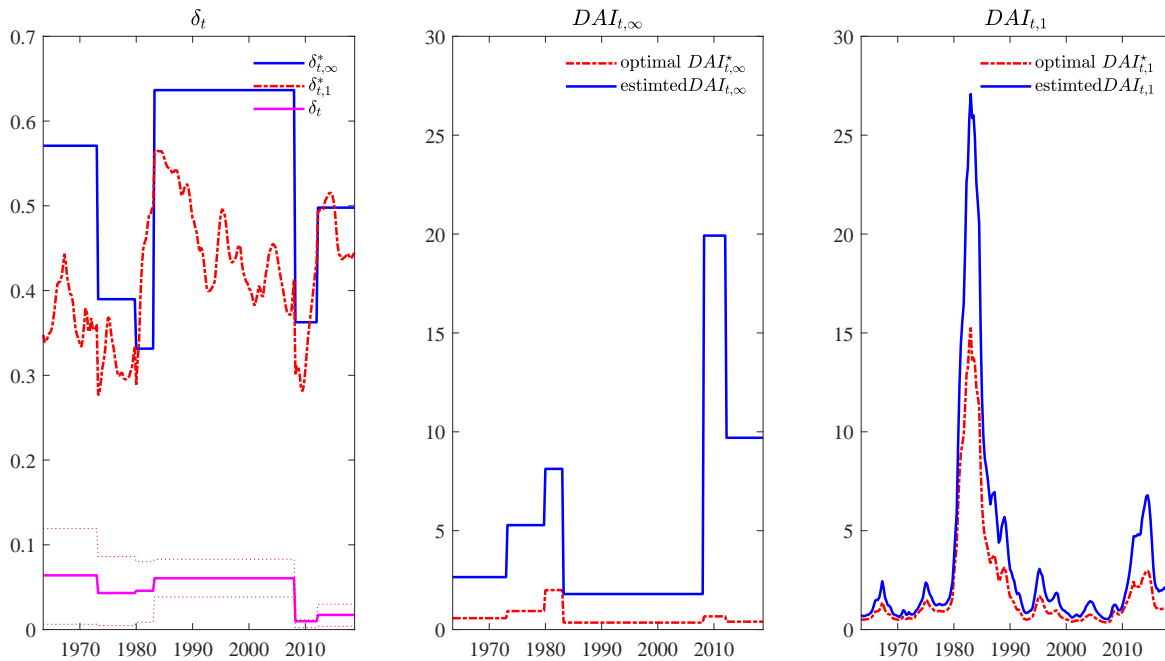
Notes: The left panel shows mean of the optimal learning rates $\delta^{\star}_{t,1}$ (solid blue) and $\delta^{\star}_{\infty}$ (dashed red) implied by the posterior distributions of the parameters, as well as the posterior mean (solid magenta) of the estimated learning rate δ with 68% credibility bands (dotted magenta) for the various monetary regimes. The middle and right panels show the posterior mean of the $DAI_{t,1}$ and DAI_{∞} evaluated at the actual learning rate (solid blue) and $DAI^{\star}_{t,1}$ and $DAI^{\star}_{\infty}$ evaluated at the respective optimal learning rates (dashed red).

Table 2.4.3: Regime dependent learning rate and shock variances

Regime	δ	σ_g^2	σ_π^2	σ_{MP}^2	σ_P^2	$\sigma_\star^2$
1972Q4	0.06 (0.01 0.12)	0.61 (0.49 0.74)	0.76 (0.58 0.94)	0.34 (0.25 0.43)	0.06 (0.04 0.09)	0.10 (0.05 0.14)
1973Q1 to 1979Q3	0.04 (0.00 0.09)	0.77 (0.59 0.95)	1.71 (1.26 2.17)	0.90 (0.60 1.19)	0.07 (0.03 0.10)	0.11 (0.04 0.16)
1979Q4 to 1982Q4	0.05 (0.01 0.08)	0.72 (0.51 0.92)	0.91 (0.61 1.21)	2.48 (1.66 3.28)	0.06 (0.03 0.08)	0.29 (0.09 0.46)
1983Q1 to 2007Q4	0.06 (0.04 0.08)	0.24 (0.21 0.28)	1.12 (0.95 1.28)	0.18 (0.14 0.22)	0.04 (0.02 0.05)	0.07 (0.04 0.10)
2008Q1 to 2011Q4	0.01 (0.00 0.01)	0.64 (0.46 0.82)	4.01 (2.91 5.08)	0.78 (0.45 1.10)	0.05 (0.03 0.08)	0.06 (0.03 0.08)
2012Q1 to 2018Q3	0.02 (0.00 0.03)	0.28 (0.21 0.34)	0.81 (0.61 1.01)	0.30 (0.21 0.39)	0.04 (0.03 0.06)	0.05 (0.03 0.08)

Notes: Reported figures are posterior means and 16% and 84% quantiles in parentheses.

value of $DAI_{t,\infty}$ should *not* be confused with poor credibility. Rather, the increase in $DAI_{t,\infty}$ signals that credibility is more vulnerable to *future* shocks. Therefore, it might be interpreted as an early warning indicator rather than a reflection of the current state. In contrast, the peak of $DAI_{t,1}$ in the early 1980s is much higher than the peak after the Financial Crisis because it takes into account that credibility gap was much smaller after the crisis.

Finally, preventing δ from becoming *too small* is more important than engineering the exact optimal value. This is due to the shape of the nonlinear mapping from the learning rate to the de-anchoring indicators. To see this, consider the de-anchoring indicators implied by the estimated parameters against the level of the indicators $DAI_{t,\infty}^\star$ and $DAI_{t,1}^\star$ that would prevail under the respective optimal learning rates, all else equal. The gap between the $DAI_{t,\infty}$ and $DAI_{t,1}$ and their optimal counterparts is small in the Great Moderation even though the deviation of δ from $\delta_\infty^\star$ is largest during this period. In contrast, the deviation of δ from $\delta_\infty^\star$ is much smaller when $DAI_{t,\infty}$ reaches its maximum. This is because the de-anchoring indicator penalizes very low learning rates over-proportionally. Thus, from a central bank's perspective, it might be more important that δ is not too small, than that it is exactly at the optimal value. In line with this conclusion, Gáti (2022) finds that a cost-push shock has almost identical effects on the typical macro variables under a *strong* and a *weak* anchoring of expectations. By contrast, the same shock has a much more adverse effect under completely unanchored expectations.

Did the 2012 announcement improve anchoring? The decline of both, $DAI_{t,\infty}$ and $DAI_{t,1}$ after the 2012 supports this hypothesis. At the same time δ recovers only marginally and cannot explain the decline in the de-anchoring indicators by itself. However, the volatility

of the temporary monetary policy shock drops after the 2012 announcement, see Table 2.4.3. This drop most likely accounts for the bulk of the improvement in the anchoring.

2.4.4 The inflation target under perfect credibility

Asymmetric information prevents credibility from being perfect even when the learning rate is at its optimal value because agents can only imperfectly disentangle between temporary monetary policy shocks and inflation target shocks. What would the estimated path of $\pi_t^\star$ look like under perfect credibility? To answer this question, the model is re-estimated under perfect information which implies $\pi_t^P = \pi_t^\star$ for all t, while maintaining time-variation in $\pi_t^\star$. Without asymmetric information $\pi_t^\star$ is the only common trend of inflation and the central bank interest rate. The implications of this restriction for the estimated path of $\pi_t^\star$ and the model fit allow for drawing conclusions about the empirical importance of imperfect credibility.[33]

Figure 2.4.4: $\pi_t^\star$ obtained under perfect credibility
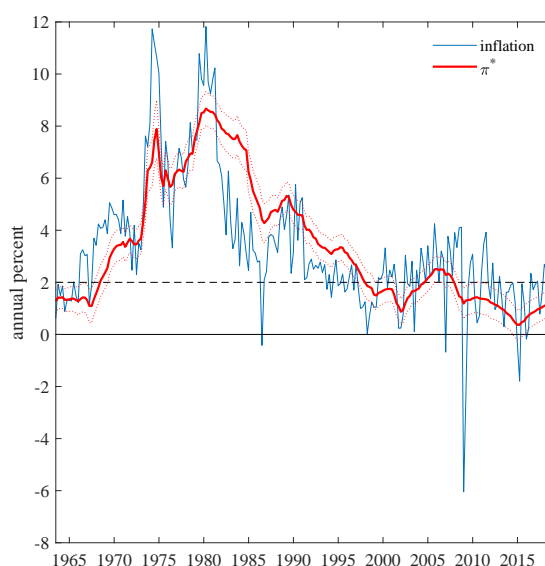
Notes: Horizontal lines indicate zero (solid) and 2% (dashed). Dotted lines are 16% and 84% posterior quantiles.

The estimate of $\pi_t^\star$, obtained under perfect credibility, reveals a counter intuitive feature: The decline of $\pi_t^\star$ during the Volcker Disinflation *lags* behind the decline of inflation, see Figure 2.4.4. This suggests that the inflation target follows the path of inflation sluggishly

---

[33]Kozicki and Tinsley (2005) only compare their asymmetric information model with a model of perfect credibility and *constant* $\pi_t^\star$, thus not allowing to draw conclusions about the statistical relevance of imperfect credibility alone. Moreover, they do not conduct a formal model comparison.

not the other way around. The literature estimating $\pi_t^\star$ under perfect credibility often finds the same counter intuitive feature, see e.g. Figure 3 of Castelnuovo et al. (2014) for a direct comparison.[34] However, this estimate of the inflation target is at odds with the view that the Volcker-Fed set out to reduce inflation already in 1979, when inflation was still very high as argued for example in Lindsey et al. (2013). Following this argument, the decline in the inflation target should *lead* the decline in inflation, as is the case in the baseline results; see Figure 2.4.2.

Moreover, a Bayesian model comparison clearly favors the baseline model against the perfect credibility model. The log marginal likelihood for the perfect credibility model is -970.21 with a numerical standard error of 0.39 against -907.93 for the baseline model; see Table 2.4.1.[35] The difference in log marginal likelihoods of approximately 62 is very strong evidence in favor of the baseline model with imperfect credibility according to the Kass and Raftery (1995) scale. Thus, the results strongly suggest that imperfect credibility is an important feature of US monetary policy throughout the postwar period.

In contrast, Del Negro and Eusepi (2011) find that rational expectations DSGE models with perefect information are preferred over asymmetric information specifications using a sample covering 1984 through 2008. However, this sample period excludes the Volcker Disinflation and the low-inflation period in the aftermath of the financial crisis, two periods where imperfect credibility appears to have been especially important according to the baseline model. Despite many other model differences, the sample period is probably an important driver of this contrasting result.

2.4.5 The 2012 announcement and SPF inflation expectations

This section explores implications of using additional information in form of the official target announcement or observed survey inflation expectations in the estimation of the model.

A 2%-prior for $\pi_t^\star$. One may argue that, since the 2012 announcement, the random walk law of motion allows $\pi_t^\star$ to wander excessively in comparison to the implicit prior belief that the inflation target should be close to 2%. To take this implicit prior into account, an explicit prior on the path of $\pi_t^\star$ can be added to the model equations. Not only is this approach preferable for its transparency instead of discarding models with 'implausible'

---

[34]Studies that estimate $\pi_t^\star$ under perfect credibility include Ireland (2007), Aruoba and Schorfheide (2011), Coibion and Gorodnichenko (2011), and Castelnuovo et al. (2014). An exception is Milani (2020) who obtains a nearly constant estimate for $\pi_t^\star$ under learning on the side of the central bank.

[35]The perfect credibility model is estimated with the same breaks as the baseline model.

trajectories for $\pi_t^\star$ *a posteriori*, but it can also help disentangling movements of $\pi_t^\star$ and π_t^P in the estimation. From a classical perspective, such a prior can be viewed as a soft restriction on the estimated path for $\pi_t^\star$. If not rejected by the data, such a restriction can lead to more economically plausible and precise estimates. The prior takes the following form

$$
\begin{aligned}
\pi_t^\star &= \pi^A + \varepsilon_t^A, \quad \varepsilon_t^A \sim N(0, s_A^2) \\
\Rightarrow \quad \pi_t^\star &\sim N(\pi^A, s_A^2) \quad \text{for} \quad t > \text{January 2012}
\end{aligned}
\tag{2.18}
$$

with $\pi^A = 2$. $\pi_t^\star$ can vary freely until January 2012 but the prior restricts time-variation thereafter. The standard deviation s_A controls the average size of deviations from the announcement π^A that are allowed under this prior. I set $s_A = 0.1$ which implies $P(1.8 < \pi_t^\star < 2.2) \approx 0.95$. Thus, most of the prior density mass of $\pi_t^\star$ is in a reasonably narrow interval around 2%.

Using survey expectations for estimation of π_t^P. Newer literature (e.g. Crump et al. 2018; Chan et al. 2018; Bańbura and van Vlodrop 2018; Del Negro et al. 2017) exploits data on long-run inflation expectations from surveys, denoted π_t^{LR}, in the estimation of the trend in inflation. Figure 2.4.5 confirms that there is a broad co-movement between 10y inflation expectations for the consumer price index (CPI) from the Philadelphia Fed's Survey of Professional Forecasters and the estimated π_t^P. To refine the estimation and relate the model based perceived target to observed inflation expectations, long-term inflation expectations could be used as a noisy measurement of π_t^P via another observation equation in the spirit of Chan et al. (2018):

$$
\begin{aligned}
\pi_t^{LR} &= d_0 + d_1 \pi_t^P + \varepsilon_t^{LR}, \quad \varepsilon_t^{LR} \sim N(0, \sigma_{LR}^2) \\
\Rightarrow \quad \pi_t^{LR} &\sim N\left(d_0 + d_1 \pi_t^P, \sigma_{LR}^2\right)
\end{aligned}
\tag{2.19}
$$

The coefficient d_0 allows for a potential bias of π_t^{LR} with respect to π_t^P and d_1 allows for different volatility of π_t^{LR} and π_t^P. A bias may arise because PCE and CPI differ on average by a constant amount. However, the dynamic properties of both series are hard to distinguish, supporting the use of CPI expectations as a measurement of the PCE based π_t^P.[36]

Table 2.4.4 compares the log marginal likelihoods of the baseline model with models using the 2%-prior in (2.18) and π_t^{LR} as in (2.19) as additional restrictions. Adding π_t^{LR}

---

[36]To relate PCE inflation to CPI long-run inflation expectations, Doh and Oksol (2018) rely on a rule of thumb that CPI is about 0.4 percentage points above PCE inflation on average.

Figure 2.4.5: π_t^p and SPF 10y CPI inflation expectations
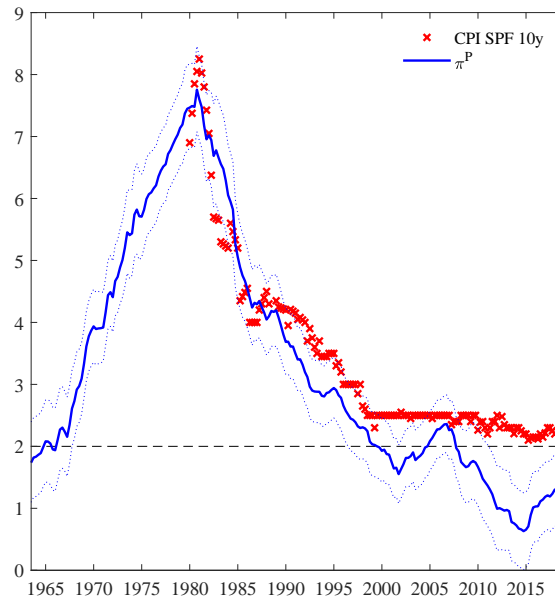
Notes: Philadelphia Fed's SPF CPI 10y inflation expectations are extended backwards using the bi-annual series of long-term from the Blue Chip and Livingston Survey, also available from the Philadelphia Fed. The final series runs from 1979Q4 to 2018Q3.

Table 2.4.4: Log marginal likelihoods of models with restrictions on π_t^P and $\pi_t^\star$

sample	Restrictions on unobserved components			
	none (baseline)	only $\pi_t^\star$	only π_t^P	$\pi_t^\star$ and π_t^P
1962–2018	−907.93 (1.22)	−910.03 (0.69)	−922.38 (0.66)	−924.15 (1.10)
1962–2007	−721.69 (0.60)	–	−737.56 (0.75)	–

Note: Reported figures are log marginal likelihoods and numerical standard error in parenthesis, see notes of Table 2.4.1. To enable an 'apples-to-apples' comparison, the likelihoods are computed based only on the macro variables y_t. According to the best-fit, break dates for all models are in January 1973, October 1979, January 1983, January 2008, and January 2012. The restrictions on π_t^P and $\pi_t^\star$ are the 2%-prior and the use of π_t^{LR} as in (2.18) and (2.19), respectively.

to the baseline model decreases the log ML by about 14 log points to −922.38 indicating a strong deterioration of the fit to the macro data. Additionally restricting $\pi_t^\star$ leads to a decrease of only 2 log points. Taking into account the numerical standard errors suggests that the deterioration in fit is not very significant. Similarly, the fit of the model with both additional restrictions is also much worse than the model only restricting $\pi_t^\star$ but not much

worse than the model that only uses π_t^{LR}. Overall, there is strong evidence against the use of π_t^{LR} to restrict π_t^P but much less compelling evidence against restricting $\pi_t^\star$ to close to 2% after 2012.

2.5 Conclusion

This paper estimates how the public forms the long-term inflation beliefs by learning from the Fed's interest rate policy, thereby contributing to a better understanding of drivers of the time-varying target credibility and the anchoring of public perceptions at the Fed's inflation target. To enable an analysis of the role of learning for the anchoring of public perceptions, I propose a de-anchoring indicator that is motivated by asymmetric information about the Fed's inflation target and derive an optimal learning rate that minimizes de-anchoring. In this model, a learning rate of zero is generally not optimal.

To apply the analysis to the US, I estimate the model on US postwar data from 1960 to 2018. To account for different monetary policy regimes, I allow for breaks in the variances and the learning rate. To estimate the model and enable Bayesian model comparison via the marginal likelihood, I derive precision based expressions for efficient state sampling and evaluation of the integrated likelihood.

Four main results emerge from the baseline estimation. First, imperfect credibility is an important feature of the joint evolution of US inflation and the Fed's interest rate policy. A model that does not allow for imperfect credibility is clearly rejected by the data. Second, the optimal learning rate varies between 0.3 and 0.65 in the US postwar period. The profile of the estimated actual learning rate largely follows the the profile of the optimal rate. However, there is a substantial level shift: The public learns much too slowly compared to the optimal rate. Third, the degree of anchoring improves after the announcement of the 2% inflation target in 2012. This improvement is mainly driven by a reduction in the volatility of temporary monetary policy shocks. The learning rate increases slightly, but this has, if at all, only a small effect on the degree of anchoring. Fourth, despite the improved anchoring, the de-anchoring indicators remain elevated compared to the Great Moderation, indicating that credibility is more vulnerable to unfavorable shocks. Credibility could be made more robust if the Fed manages to increase the weight of monetary policy in agent's belief formation process.

Finally, this model shows that a decline in the learning rate can also deteriorate the degree of anchoring of public perceptions to the actual inflation target. In contrast, a learning rate of zero maximizes anchoring in models where agents learn from inflation

surprises instead of monetary policy. For future research it would be interesting to pin down the relative importance of these two sources of learning, especially since central banks around the world are confronted with above target inflation again since 2021.

Appendix

2.A Analytical Results

Proof 1 (Proof of Proposition 1) *Let $\pi_t^\dagger = \pi_t^P - \pi_t^\star$ denote the credibility gap and its law of motion*

$$\pi_t^\dagger = \alpha \pi_{t-1}^\dagger + u_t, \quad u_t \sim N(0, \sigma_{u,t}^2)$$

$$\text{with} \quad \alpha = (1 - \delta_t \phi_\pi)$$

$$u_t = \varepsilon_t^P - \delta_t \varepsilon_t^{MP} - (1 - \delta \phi_\pi) \varepsilon_t^\star$$

$$\text{and} \quad \sigma_{u,t}^2 = \sigma_P^2 + \delta_t^2 \sigma_{MP}^2 + (1 - \delta \phi_\pi)^2 \sigma_\star^2$$

Iterating $\pi_t^\dagger$ forward yields

$$\pi_{t+s}^\dagger = \alpha^s \pi_t^\dagger + \sum_{i=1}^{s} \alpha^{s-i} u_{t+s-i}.$$

Using this expression, the $DAI_{t,s} = E\left[\left(\pi_{t+s}^\dagger \right)^2 \middle| \mathcal{I}_t \right]$ can be written as

$$
\begin{aligned}
DAI_{t,s} &= \alpha^{2s}(\pi_t^\dagger)^2 + \sum_{i=1}^{s} \alpha^{2(s-i)} \sigma_u^2 \\
&= \alpha^{2s}(\pi_t^\dagger)^2 + \frac{1 - \alpha^{2s}}{1 - \alpha^2} \sigma_u^2 \\
&= (\pi_t^\dagger)^2 (1 - \delta \phi_\pi)^{2s} + \left(1 - (1 - \delta \phi_\pi)^{2s} \right) \frac{\left(\delta^2 \sigma_{MP}^2 + \sigma_*^2 (1 - \delta \phi_\pi)^2 + \sigma_P^2 \right)}{1 - (1 - \delta \phi_\pi)^2}
\end{aligned}
$$

Proof 2 (Proof of Proposition 2) *For $s \to \infty$*

$$\frac{\partial \, \mathrm{DAI}_\infty}{\partial \delta} = \frac{2 \left(\delta^2 \sigma_{MP}^2 + (\sigma_*^2 + \sigma_P^2)(\delta \phi_\pi - 1) \right)}{\delta^2 \phi_\pi (\delta \phi_\pi - 2)^2}$$

with $\frac{\partial \, \mathrm{DAI}_\infty}{\partial \delta}\Big|_{\delta = \delta_\infty^\star} = 0.$ *It remains to show that* $\frac{\partial^2 \, \mathrm{DAI}}{\partial \delta^2} > 0,$ *where*

$$\frac{\partial^2 \, \mathrm{DAI}_\infty}{\partial \delta^2} = -\frac{2 \left(2\delta^3 \phi_\pi \sigma_{MP}^2 + (\sigma_*^2 + \sigma_P^2)(3\delta^2 \phi_\pi^2 - 6\delta\phi_\pi + 4)\right)}{\delta^3 \phi_\pi (\delta\phi_\pi - 2)^3}.$$

The denominator of $\frac{\partial^2 \, \mathrm{DAI}_\infty}{\partial \delta^2}$ *is negative for the admissible parameters space* $\delta \in (0, \frac{1}{\phi_\pi})$. *For admissible values of* δ, *the nominator is positive if* $3(\delta\phi_\pi)^2 - 6\delta\phi_\pi + 4 > 0$. *This quadratic form is bounded from below by unity for* $\delta\phi_\pi = 1$ *at the upper bound of the admissible parameter space. Therefore,* $\frac{\partial^2 \, \mathrm{DAI}_t}{\partial \delta^2}$ *is negative everywhere in the admissible parameter space.*

   For $s = 1$

$$\frac{\partial \, \mathrm{DAI}_{t,1}}{\partial \delta} = 2(\pi_t^P - \pi_t^\star)^2 \phi_\pi (\delta\phi_\pi - 1) + 2\delta\sigma_{MP}^2 + 2\phi_\pi \sigma_*^2 (\delta\phi_\pi - 1)$$

with $\frac{\partial \, \mathrm{DAI}_{t,1}}{\partial \delta}\Big|_{\delta = \delta_{t,1}^\star} = 0.$ *Furthermore, we have*

$$\frac{\partial^2 \, \mathrm{DAI}_{t,1}}{\partial \delta^2} = 2(\pi_t^P - \pi_t^\star)^2 \phi_\pi^2 + 2\sigma_{MP}^2 + 2\phi_\pi^2 \sigma_*^2 > 0.$$

2.B   Model matrices

The full model is given by

$$y_t = \tilde{\gamma}_y + \Gamma_0 \tau_t + \Gamma_1 \tau_{t-1} + c_t \tag{2.20}$$

$$Ac_t = B_1 c_{t-1} + \ldots + B_p c_{t-p} + \lambda_0 \tau_t + \ldots + \lambda_q \tau_{t-q} + e_t, \quad e_t \sim N(0, \Sigma_t) \tag{2.21}$$

$$\tau_t = F_t \tau_{t-1} + J_t e_t + O_t u_t, \quad u_t \sim N(0, \Omega_t). \tag{2.22}$$

with diagonal variance matrices $\Sigma_t = \mathrm{diag}\left(\sigma_{g,t}^2, \sigma_{\pi,t}^2, \sigma_{MP,t}^2\right)$ and $\Omega_t = \mathrm{diag}\left(\sigma_{P,t}^2, \sigma_{\star,t}^2\right)$, and

$$\tilde{\gamma}_y = \begin{bmatrix} \bar{g} \\ 0 \\ rr \end{bmatrix} \qquad \Gamma_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \Gamma_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

B. Model matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ a_{2,1} & 1 & 0 \\ -\phi_y & -\frac{\phi_\pi}{4} & 1 \end{bmatrix} \qquad \lambda_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\phi_\pi}{4} & -\phi_\pi \end{bmatrix} \qquad \lambda_i = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\phi_\pi}{4} & 0 \end{bmatrix} \quad \text{for} \quad i = 1, ..., 3$$

$$F_t = \begin{bmatrix} 1 - \phi_\pi \delta_t & \phi_\pi \delta_t \\ 0 & 1 \end{bmatrix} \qquad O_t = \begin{bmatrix} 1 & \phi_\pi \delta_t \\ 0 & 1 \end{bmatrix} \qquad J_t = \begin{bmatrix} 0 & 0 & -\delta_t \\ 0 & 0 & 0 \end{bmatrix}.$$

Stacking up over T yields

$$y = \gamma_y + \gamma_\tau + \Gamma\tau + c \tag{2.23}$$

$$H_{A,\beta}c = \Lambda\tau + \gamma_c + e, \qquad\qquad e \sim N(0, \Sigma) \tag{2.24}$$

$$H_F\tau = \alpha_\tau + Je + Ou, \qquad\qquad u \sim N(0, \Omega) \tag{2.25}$$

with

$$\Gamma = \text{blockdiag}\left([\Gamma_0', \Gamma_1']'\right), \qquad O = \text{blockdiag}\left(O_1, ..., O_T\right), \qquad J = \text{blockdiag}\left(J_1, ..., J_T\right)$$

$$\Sigma = \text{blockdiag}\left(\Sigma_1, ..., \Sigma_T\right), \qquad \Omega = \text{blockdiag}\left(\Omega_1, ..., \Omega_T\right)$$

The impact of initial values is collected in

$$\gamma_y = \left(\mathbf{1}_T \otimes \tilde{\gamma}_y\right)$$

$$\gamma_c = \left[\left(\begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_{q-1} & \lambda_q \\ \lambda_2 & \lambda_3 & \dots & \lambda_q & 0 \\ \vdots & & & & \vdots \\ \lambda_q & 0 & \dots & & 0 \end{bmatrix} \begin{bmatrix} \tau_0 \\ \tau_{-1} \\ \vdots \\ \tau_{-q+1} \end{bmatrix}\right)' \quad 0_{1\times(T-q)r} \right]' \tag{2.26}$$

$$\gamma_\tau = [\Gamma_1'\tau_0, \, 0_{1\times n(T-1)}]'$$

$$\alpha_\tau = [(F_1\tau_0)', 0_{1\times T(r-1)}]'$$

71

and the large coefficient matrices are

$$
\Lambda = \begin{bmatrix} \lambda_0 & 0 & \dots & & & & 0 \\ \lambda_1 & \lambda_0 & 0 & \dots & & & 0 \\ & \ddots & \ddots & & \ddots & & \vdots \\ \lambda_q & \dots & \lambda_1 & \lambda_0 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \ddots & & \\ \vdots & & & & & & 0 \\ 0 & \dots & 0 & \lambda_q & \dots & \lambda_1 & \lambda_0 \end{bmatrix} \qquad H_F = \begin{bmatrix} I_r & 0 & \dots & & & & 0 \\ -F_2 & I_r & 0 & \dots & & & 0 \\ & \ddots & \ddots & & \ddots & & \vdots \\ 0 & \dots & -F_s & I_r & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \ddots & & \\ \vdots & & & & & & 0 \\ 0 & \dots & 0 & 0 & \dots & -F_T & I_r \end{bmatrix}
$$

$$
H_{A,\beta} = \begin{bmatrix} A & 0 & 0 & & & & \dots & 0 \\ -B_1 & A & 0 & & & & \dots & 0 \\ -B_2 & -B_1 & A & 0 & & & \dots & 0 \\ \vdots & & \ddots & \ddots & & & & \vdots \\ -B_p & \dots & & -B_1 & A & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \ddots & \ddots & & \vdots \\ \vdots & & & & & A & 0 \\ 0 & \dots & 0 & -B_p & \dots & & -B_1 & A \end{bmatrix}
$$

$$(2.27)$$

2.C Bayesian model comparison and estimation of the MCUC

This sections briefly introduces the notation and Bayesian concepts that are useful for understanding the derivation in the next section and the results from model comparison. It draws heavily on Chan and Eisenstat (2015), Chan and Grant (2015), and Chan and Grant (2016).

Let M_k denote model k and y be the data vector. Bayesian model comparison is conducted by a comparison of the marginal data densities or marginal likelihoods (ML) $p(y|M_k)$ and $p(y|M_j)$ of models k and j. Akin to likelihood ratio tests, the evidence in favor of model k over model j is given by the ratio of the marginal likelihoods of the two models, the so-called *Bayes Factor* BF_{kj}:

$$
BF_{kj} = \frac{p(y|M_k)}{p(y|M_j)} \tag{2.28}
$$

The ML of a model is be obtained from prior and posterior densities of the model. To simplify the notation, I omit the the explicit dependence on model M_k and let θ collect all model parameters and let $p(\theta)$ be the prior density. Then $p(y|\theta)$ is referred to as the *observed-data* likelihood that is implied by the model equations. The marginal likelihood $p(y)$ of a model is obtained by integrating out the unknown parameters from the observed data likelihood via

$$p(y) = \int p(y|\theta)p(\theta)d\theta. \tag{2.29}$$

In most cases, this integral has to be solved numerically, which requires evaluation of $p(y|\theta)$. However, for state space models like the MCUC that contain a vector of latent state variables τ, the observed-data likelihood $p(y|\theta)$ cannot readily be evaluated analytically. Instead, only the likelihood $p(y|\tau, \theta)$ *conditional* on the latent states τ can be evaluated directly. To evaluate the observed data, the states τ have to be integrated out via

$$p(y|\theta) = \int p(y, \tau|\theta)d\tau = \int p(y|\tau, \theta)p(\tau|\theta)d\tau. \tag{2.30}$$

where $p(y, \tau|\theta)$ is called the *complete* data likelihood to distinguish it from the conditional likelihood $p(y|\tau, \theta)$. To solve this integral, analytical expressions are available for many linear state space models including the univariate correlated unobserved components model of Grant and Chan (2017). However, to the best of my knowledge, there is no analytical expression for $p(y|\theta)$ that can suitable for the MCUC. Therefore, in the next section I generalize the results in Grant and Chan (2017) to a multivariate setting. Being able to evaluate $p(y|\theta)$ analytically enables fast Bayesian model comparison, and Bayesian and maximum likelihood estimation.[37]

2.C.1   The observed data density of the MCUC

This section outlines the derivation of $p(y|\theta)$ and its components. For convenience, I repeat the laws of motions for the vectorized states $\tau = [\tau_1', ..., \tau_T']'$ and $c = [c_1', ..., c_T']'$, and the

---

[37]A less computationally efficient solution would be the Kalman Filter.

observation equation for $y = [y_1', ..., y_T']'$:

$$y = \gamma_y + \gamma_\tau + \Gamma\tau + c$$
$$H_{A,\beta}c = \Lambda\tau + \gamma_c + e, \qquad\qquad e \sim N(0, \Sigma)$$
$$H_F\tau = \alpha_\tau + Je + Ou, \qquad\qquad u \sim N(0, \Omega)$$

Definitions of the matrices $H_{A,\beta}$, H_F, and the terms for the initial values are in Appendix 2.B. Note that $\det(H_{A,\beta}) = \det(H_F) = 1$. Inverting $H_{A,\beta}$ and H_F and absorbing the term involving τ from the equation for the cycles c into the observation equation yields

$$y = \gamma_y + \gamma_\tau + X_{y,\tau}\tau + \tilde{c} \tag{2.31}$$
$$\tilde{c} = m_{\tilde{c}} + H_{A,\beta}^{-1}e \tag{2.32}$$
$$\tau = H_F^{-1}\alpha_\tau + H_F^{-1}Je + H_F^{-1}Ou \tag{2.33}$$

with

$$m_{\tilde{c}} = H_{A,\beta}^{-1}\gamma_c, \qquad\qquad m_\tau = H_F^{-1}\tilde{\alpha}_{\tilde{\tau}} \qquad\qquad X_{y,\tau} = \Gamma + H_{A,\beta}^{-1}\Lambda.$$

The joint distribution of $\tilde{c}$ and τ is implied by model equations (2.33) and (2.32), and reads

$$\begin{pmatrix} \tau \\ \tilde{c} \end{pmatrix} \sim N\left(\begin{pmatrix} m_\tau \\ m_{\tilde{c}} \end{pmatrix}, \begin{pmatrix} H_F^{-1}\Sigma_\tau(H_F^{-1})' & H_F^{-1}\Sigma_{\tau,\tilde{c}}(H_{A,\beta}^{-1})' \\ H_{A,\beta}^{-1}\Sigma_{\tau,\tilde{c}}'(H_F^{-1})' & H_{A,\beta}^{-1}\Sigma_{\tilde{c}}(H_{A,\beta}^{-1})' \end{pmatrix} \right)$$

$$\text{with} \quad \Sigma_\tau = O\Omega O' + J\Sigma J'$$
$$\Sigma_{\tilde{c}} = \Sigma$$
$$\Sigma_{\tau,\tilde{c}} = J\Sigma$$

This implies the marginal distribution, for τ with precision matrix K_τ as

$$\tau|\theta \sim N(m_\tau, K_\tau^{-1}) \quad \text{with} \quad K_\tau = H_F'\Sigma_\tau^{-1}H_F \tag{2.34}$$

and corresponding density function

$$p(\tau|\theta) = (2\pi)^{-\frac{rT}{2}} \det(K_\tau^{-1})^{-\frac{1}{2}} e^{-\frac{1}{2}(\tau - m_\tau)'K_\tau(\tau - m_\tau)}. \tag{2.35}$$

From the formulas for conditional normal distributions (see e.g. Kroese et al. 2014, Chapter 3.6), the conditional distribution of $(\tilde{c} \,|\, \tau, \theta)$ is

$$
\begin{aligned}
\tilde{c} \,|\, \tau, \theta &\sim N\left(m_{\tilde{c}} + H_{A,\beta}^{-1} B H_F(\tau - m_\tau),\ K_{\tilde{c}|\tau}^{-1}\right) \\
\text{with} \quad K_{\tilde{c}|\tau}^{-1} &= H_{A,\beta}^{-1} P H_{A,\beta}^{-1}{}' \\
B &= \Sigma_{\tau,\tilde{c}}' \Sigma_\tau^{-1} = \Sigma J'(O\Omega O' + J\Sigma J')^{-1} \\
P &= \Sigma - \Sigma J'(O\Omega O' + J\Sigma J')^{-1} J\Sigma
\end{aligned}
\tag{2.36}
$$

In the case of constant variances and learning gain δ, the matrices O, J, Ω and Σ have a Kronecker structure and the expressions for B and P could be further simplified. To handle breaks in variances I stick to the more general expressions in this derivation. The distribution in (2.36) and the observation equation (2.31) together imply the condition distribution $(y \,|\, \tau, \theta)$

$$
\begin{aligned}
y \,|\, \tau, \theta &\sim N\left(m_{y,\tau} - H_{A,\beta}^{-1} B H_F m_\tau + X\tau,\ K_{y|\tau}^{-1}\right) \\
\text{with} \quad m_{y,\tau} &= \gamma_y + H_{A,\beta}^{-1}\gamma_c + \gamma_\tau \\
X &= H_{A,\beta}^{-1} B H_F + \Gamma + H_{A,\beta}^{-1}\Lambda \\
K_{y|\tau} &= K_{\tilde{c}|\tau} = H_{A,\beta}' P^{-1} H_{A,\beta}
\end{aligned}
\tag{2.37}
$$

with the corresponding conditional data density

$$
p(y\,|\,\tau,\theta) = (2\pi)^{-\frac{nT}{2}} \det(K_{y|\tau}^{-1})^{-\frac{1}{2}} e^{-\frac{1}{2}(y-m_{y,\tau}-X\tau)' K_{y|\tau}(y-m_{y,\tau}-X\tau)}.
\tag{2.38}
$$

The complete data density $p(y, \tau\,|\,\theta)$ is obtained as the product of the two densities in $p(\tau\,|\,\theta)$ in (2.35) and $p(y\,|\,\tau,\theta)$ in (2.38). Applying the steps in the appendix of Grant and Chan (2017) to solve the integral in (2.30) yields the observed data density $p(y\,|\,\theta)$. To that end, define $c_1 = (2\pi)^{-\frac{(n+r)T}{2}} \det(K_{y|\tau}^{-1})^{-\frac{1}{2}} \det(K_\tau^{-1})^{-\frac{1}{2}}$ and rewrite

$$
\begin{aligned}
p(y\,|\,\theta) &= \int p(y,\tau|\theta)d\tau = \int p(y|\tau,\theta)p(\tau|\theta)d\tau \\
&= c_1 \int e^{-\frac{1}{2}\left[(y-m_{y,\tau}-X\tau)' K_{y|\tau}(y-m_{y,\tau}-X\tau)+(\tau-m_\tau)' K_\tau(\tau-m_\tau)\right]}d\tau
\end{aligned}
$$

After some algebra, the observed data density can be written as

$$
= (2\pi)^{-\frac{nT}{2}} \det(K_{y|\tau}^{-1})^{-\frac{1}{2}} \det(K_\tau^{-1})^{-\frac{1}{2}} \det(P_\tau^{-1})^{\frac{1}{2}} e^{-\frac{1}{2}\left[(y-m_{y,\tau})' K_{y|\tau}(y-m_{y,\tau})+m_\tau' K_\tau m_\tau - d_\tau' P_\tau^{-1} d_\tau\right]}.
$$

with $P_\tau = X'K_{y|\tau}X + K_\tau$ and $d_\tau = X'K_{y|\tau}(y - m_{y,\tau}) + K_\tau m_\tau$. This expression does not depend on the states *tau* anymore and thus, be evaluated directly. I employ this expression in the estimation of the marginal data density of the MCUC using the cross-entropy method of Chan and Eisenstat (2015).

2.C.2 Drawing the states τ

A by product of this derivation is an analytical expression for $p(y, \tau | \theta)$. Bayes' formula shows that the full conditional posterior $p(\tau | y, \theta)$ is proportional to $p(y, \tau | \theta)$.

$$p(\tau | y, \theta) = \frac{p(y | \tau, \theta)p(\tau | \theta)}{p(y)}$$
$$\propto p(y | \tau, \theta)p(\tau | \theta) = p(y, \tau | \theta)$$

Then, from the above derivation it follows that

$$p(\tau | y, \theta) \propto e^{-\frac{1}{2}\left[(\tau - P_\tau^{-1}d_\tau)'P_\tau(\tau - P_\tau^{-1}d_\tau)\right]}.$$

This is the kernel of the multivariate normal distribution $N(\hat{\tau}, P_\tau^{-1})$ with $\hat{\tau} = P_\tau^{-1}d_\tau$. The precision sampler of Chan and Jeliazkov (2009) can be used to generate draws of τ in a Gibbs sampler that simulates the posterior of the entire model.

2.D Gibbs sampler for the MCUC with breaks in variances and the learning rate

This section presents the details of the Gibbs sampler for the baseline model with breaks in the shock variances and the learning rate. To that end, let

$$\theta = \left\{A, \beta, \delta, \bar{g}, rr, d, \sigma_g^2, \sigma_\pi^2, \sigma_{MP}^2, \sigma_P^2, \sigma_\star^2, \sigma_{LR}^2\right\}$$

collect all parameters that make up the model matrices and let θ_{-i} be all parameters except parameter set i. The Gibbs sampler to estimate the model in (2.23) to (2.25) will consists of iteration of the following steps:

1. Sample τ jointly.

2. Sample the free parameters in γ_y and $\tilde{\Gamma}$.

3. Sample the free parameters in β and λ equation by equation, subject to stability of the cycles.

4. Sample δ, the parameter in the trend equation.

5. Sample the shock variances Σ and Ω.

6. Sample the p initial values τ_0.

Each parameter block is sampled from its full conditional posterior density. The following presents the details of the densities for each step.

1. Sampling the states τ jointly. A by product of the derivation of the marginal data density of the MCUC in Appendix 2.C.1 is an analytical expression for $p(y, \tau \mid \theta)$. Bayes' formula shows that the full conditional posterior $p(\tau \mid y, \theta)$ is proportional to $p(y, \tau \mid \theta)$.

$$
\begin{aligned}
p(\tau \mid y, \theta) &= \frac{p(y \mid \tau, \theta) p(\tau \mid \theta)}{p(y)} \\
&\propto p(y \mid \tau, \theta) p(\tau \mid \theta) = p(y, \tau \mid \theta)
\end{aligned}
$$

It follows that

$$
p(\tau \mid y, \theta) \propto e^{-\frac{1}{2}\left[(\tau - P_\tau^{-1} d_\tau)' P_\tau (\tau - P_\tau^{-1} d_\tau) \right]}.
$$

This is is the kernel of the multivariate normal distribution $N(\hat{\tau}, P_\tau^{-1})$ with $\hat{\tau} = P_\tau^{-1} d_\tau$. The precision sampler of Chan and Jeliazkov (2009) can be used to generate draws of τ in a Gibbs sampler that simulates the posterior of the entire model.

2. Sampling the parameters in the observation equation.
To account for linear restrictions on free the parameters in the constants $\tilde{\gamma}_y$ and factor loadings $\tilde{\Gamma}$, jointly denoted as $\bar{\gamma} = \mathrm{vec}([\tilde{\gamma}_y, \Gamma_0, \Gamma_1]')$, let the $\bar{\gamma}_f$ collect the free elements that are related to $\bar{\gamma}$ via

$$
\bar{\gamma} = R_\gamma \bar{\gamma}_f + r_\gamma.
$$

In the present model, $\bar{\gamma}_f = [\bar{g}, rr]'$. The posterior of γ_f is given by

$$
p(\gamma_f \mid y, \tau, \theta_{-\gamma_f}) \propto p(y \mid \tau, \theta) p(\gamma_f)
$$

γ_f appears in the observation equation (2.23). Plugging in (2.24) and collecting terms to write in terms of the correlated errors e yields

$$y = \gamma_y + \gamma_\tau + (\Gamma + H_{A,\beta}^{-1}\Lambda)\tau + H_{A,\beta}^{-1}e.$$

Define $\gamma_\tau + \gamma_y + \Gamma\tau = \tilde{X}_\gamma\bar{\gamma}$ to factor out $\bar{\gamma}$ yields

$$y = H_{A,\beta}^{-1}\Lambda\tau + \tilde{X}_{y\gamma}\bar{\gamma} + \tilde{c}$$

$$\text{with} \qquad \tilde{X}_\gamma' = [x_{y\tau,1}, ..., x_{y\tau,T}]' \quad \text{and} \quad x_{y\tau,t}' = I_n \otimes [1\,\tau_t'\,\tau_{t-1}'].$$

Inserting the linear restrictions for $\bar{\gamma}$ implies the following conditional likelihood via (2.37) in terms of the free parameters $\bar{\gamma}_f$

$$y|\,\theta,\tau \sim N(m_{y,\gamma} + X_{y\gamma}\bar{\gamma}_f, K_{y|,\tau}^{-1})$$

$$\text{with} \qquad m_{y,\gamma} = H_{A,\beta}^{-1}(BH_F\tau - B\alpha_\tau) + \tilde{X}_\gamma r_\gamma + H_{A,\beta}^{-1}(\Lambda\tau + \gamma_c)$$

$$\text{and} \qquad X_\gamma = \tilde{X}_\gamma R_\gamma$$

Combining this conditional likelihood with a normal prior $\gamma_f \sim N(a_\gamma, V_\gamma)$ yields the posterior via standard regression results:

$$\gamma_f|\,y,\theta,\tau \sim N(\hat{\gamma}_f, K_\gamma^{-1})$$

$$\text{with} \qquad K_\gamma = V_\gamma^{-1} + X_\gamma'K_{y|\tau}X_\gamma \qquad (2.39)$$

$$\text{and} \qquad \hat{\gamma}_f = K_\gamma^{-1}\left(V_\gamma^{-1}a_\gamma + (H_{A,\beta}X_\gamma)'P^{-1}y_\gamma^\star\right)$$

where $y_\gamma^\star = H_{A,\beta}(y - m_{y,\gamma}) = H_{A,\beta}(y - \tilde{X}_\gamma r_\gamma) - (BH_F + \Lambda)\tau - \gamma_c - B\alpha_\tau$.

3. Sampling the free coefficients in A, $B_1, ..., B_p$ and Λ

The coefficients in the cycles and τ are connected via linear cross equation restrictions due to the learning mechanism. Therefore, they must be sampled jointly. Since A is recursive, the equation for in c can be written as a system of equations by bringing A to the right hand side as follows:

$$c_t = -(A - I_n)c_t + B_1c_{t-1} + ... + B_pc_{t-p} + \lambda_0\tau_t + ... + \lambda_q\tau_{t-q} + e_t, \quad e_t \sim N(0, \Sigma). \quad (2.40)$$

Due to the recursive structure in A, the term $-(A - I_n)c_t$ does not introduce dependence of c_{it} to itself. Next, let β collect the free elements in vectorized form. They appear in the

vectorized equations for c in for following way

$$c = X_c R_{c,\beta} \beta + e \tag{2.41}$$

with $\quad R_{c,\beta} \beta = \begin{bmatrix} \text{vec}(-(A - I_n))' & \text{vec}([B_1 \, ... \, B_p])' & \text{vec}([\lambda_0 \, ... \, \lambda_q])' \end{bmatrix}'$ $\tag{2.42}$

and with a total of $n_A = \frac{n(n-1)}{2}$ parameters in the recursive A matrix and $(i - 1)$ parameters in the ith row. Additionally, equation i contains np parameters in the B matrices each qr parameters in the λs. In total there are $n_B = n^2 p$ and $n_\lambda = nqr$ parameters in the B's λs. However, these parameters are subject to restrictions such that the number of free elements in β in equation i is denoted k_i. Let the total number of free parameters be $K = \sum_{i=1}^{n} k_i$. Since only the monetary policy equation is restricted and only the λs in this equation contain non-zero elements, k_i equals the total number of parameters per equation for the first two equations. Thus, we have $k_i = np$ for $i < 3$. Furthermore, $k_3 = 1$ because the only free element on lagged variables in the third equation is the interest rate smoothing parameter. Hence, the total number of free parameters is $K_\beta = n_A + K$ and $R_{c,\beta}$ can be written as follows

$$\underset{n(np+\frac{(n-1)}{2}+qr)\times K_\beta}{R_{c,\beta}} = \begin{bmatrix} & 0_{n_A \times K} \\ \underset{n(np+\frac{(n-1)}{2}+qr)\times n_A}{R_{A,\beta}} & R_{B,\beta} \\ & 0_{n_\lambda \times K} \end{bmatrix} \tag{2.43}$$

$$\underset{n^2 p \times K}{R_{B,\beta}} = \begin{bmatrix} I_{np} & 0_{np \times k_2} & 0_{np \times k_3} \\ 0_{np \times k1} & I_{np} & 0_{np \times k_3} \\ 0_{np \times k_1} & 0_{np \times k_2} & \underset{np \times 1}{R_{B,\rho}} \end{bmatrix} \tag{2.44}$$

The only non-zero element in $\underset{np \times 1}{R_{B,\rho}}$ is a one in column n. The only linear restriction concerns the monetary policy reaction coefficient to inflation ϕ_π. ϕ_π appears in A as well as in the B's ans λ's. Therefore

$$\underset{n(np+\frac{(n-1)}{2}+qr)\times n_A}{R_{A,\beta}} = \begin{bmatrix} I_{n_A-1} \\ 0_{n^2 p \times n_A} & R_{\phi_\pi,\beta} \\ 0_{nqr \times n_A} \end{bmatrix}$$

with the column vector o $R_{\phi_\pi,\beta}$ of size $n(np + \frac{(n-1)}{2} + qr) \times 1$. $R_{\phi_\pi,\beta}$ has 0.25 at positions n_A and $n_A + k_1 + k_2 + 2 + (j - 1)n$ for $j = 1, .., p - 1$ for the contemporaneous and lagged reaction to the inflation gap (i.e. 4-period average), at positions $n_A + K + i_{\pi^P} + (j - 1)r$

for the reaction to the 4-period average of π^P and, finally, -1 at position $n_A + K + r$ for the reaction to $\pi_t^\star$, where $\pi_t^\star$ is ordered last within the trends τ).

ϕ_π, which is part of the coefficient vector β, also appears in the law of motion for τ. Due to the upper triangular structure of O_t it hold that $O_t^{-1} F_t = \begin{bmatrix} 1 & -\phi_\pi \delta_t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 - \phi_\pi \delta_t & \phi_\pi \delta_t \\ 0 & 1 \end{bmatrix} =$
$\begin{bmatrix} 1 - \phi_\pi \delta_t & 0 \\ 0 & 1 \end{bmatrix}$ so the law of motion can be rewritten as

$$O_t^{-1} \tau_t = \tau_{t-1} + J_t e_t + u_t$$
$$\tau_t = -(O_t - I)\tau_{t-1} + J_t e_t + u_t$$

For deriving the posterior of β rewrite the law of motion for the trends as

$$\tau = X_\tau (R_{\tau,\beta}\beta + r_{\tau,\beta}) + (I_t \otimes J)e + (I_T \otimes O)u$$

with
$$X_\tau = [x'_{\tau,1}, ..., x'_{\tau,T}]$$
and
$$x_{\tau,t} = I_2 \otimes [\tau'_t, \tau'_{t-1}]$$

The full conditional posterior density of β denoted $p(\beta | y, \tau, \theta_{-\beta})$ is, thus, obtained by

$$p(\beta | y, \tau, \theta_{-\beta}) \propto p(c, \tau | \beta, \theta_{-\beta}) p(\beta)$$
$$= p(c | \tau, \beta, \theta_{-\beta}) p(\tau | \beta, \theta_{-\beta}) p(\beta).$$

I proceed by first deriving the marginal distribution of τ, denoted $p(\tau | \beta, \theta_{-\beta})$, and then the conditional distribution of c, denoted $p(c | \tau, \beta, \theta_{-\beta})$, from joint distribution of τ and c. The joint distribution of c and τ is given by

$$\begin{pmatrix} \tau \\ c \end{pmatrix} \sim N \left(\begin{pmatrix} X_\tau (R_{\tau,\beta}\beta + r_{\tau,\beta}) \\ X_c R_{c,\beta}\beta \end{pmatrix}, \begin{pmatrix} \Sigma_\tau & \Sigma_{\tau,\tilde{c}} \\ \Sigma_{\tau,\tilde{c}} & \Sigma_{\tilde{c}} \end{pmatrix} \right)$$

with
$$\Sigma_\tau = O\Omega O' + J\Sigma J'$$
$$\Sigma_{\tilde{c}} = \Sigma$$
$$\Sigma_{\tau,\tilde{c}} = J\Sigma$$

Therefore, the marginal distribution of τ is

$$\tau \sim N \left(X_\tau (R_{\tau,\beta}\beta + r_{\tau,\beta}), \Sigma_\tau \right).$$

Applying the formulas for conditional normal distributions, the distribution of c conditional on τ is found as

$$c\,|\,\tau,\theta \sim N\left(X_\beta\beta + m_{c,\beta},\ K_{c|\tau}^{-1}\right)$$

$$\text{with} \qquad X_\beta = X_c R_c - BX_\tau R_\tau$$

$$m_{c,\beta} = B(\tau - X_\tau r_{\tau,\beta})$$

$$B = \Sigma'_{\tau,\tilde{c}}\Sigma_\tau^{-1}$$

$$K_{c|\tau}^{-1} = \Sigma_{\tilde{c}} - \Sigma'_{\tau,\tilde{c}}\Sigma_\tau^{-1}\Sigma_{\tau,\tilde{c}}$$

$$= P$$

Combining the densities with the with the normal prior $\beta \sim N(\beta_0, V_\beta)$ via Bayes' rule yields the following posterior:

$$\beta|\,y,\tau,\theta_{-\beta} \sim N(\hat{\beta}, K_\beta^{-1})$$

$$\text{with} \qquad K_\beta = V_\beta^{-1} + X'_\beta K_{c|\tau}X_\beta + (X_\tau R_{\tau,\beta})'K_\tau X_\tau R_{\tau,\beta} \qquad (2.45)$$

$$\text{and} \qquad \hat{\beta} = K_\beta^{-1}\left(V_\beta^{-1}\beta_0 + X'_\beta K_{c|\tau}(c - m_{c,\beta}) + (X_\tau R_{\tau,\beta})'K_\tau(\tau - X_\tau r_{\tau,\beta})\right)$$

4. Sampling δ

The learning rate δ_t appears only in the law of motion for the perceived target. One complication arises because $\delta_t > 0$ for each regime. To account for this inequality restriction, the parameter can be sampled with a Griddy Gibbs step. The Griddy Gibbs step requires a closed interval. Therefore, to apply it to the sampling of δ_t, I implement an upperbound $\delta_t < \delta_{\text{ub}}$ that is large enough that it does not constrain the estimate for δ in the empirical applcation. The Griddy Gibbs step also requires the full conditional posterior of δ_t. To that end, rewrite the law of motion for π_t^P

$$\pi_t^P = (1 - \delta_t\phi_\pi)\pi_{t-1}^P + \delta_t\phi_\pi\pi_t^\star - \delta_t\varepsilon_t^{MP} + \varepsilon_t^P$$

$$\text{from (2.6) as} \qquad \pi_t^P = X_\delta(R_\delta\delta + r_\delta) + \varepsilon_t^P$$

$$\text{with} \qquad X_\delta = \left[[\pi_0^P, \pi_{[1:T-1]}^P\,']',\ \pi^\star,\ \varepsilon_t^{MP}\right]$$

where $\pi_{[1:T-1]}^P$ are the elements of π^P from $t = 1$ to $t = T-1$. The vectors $R_\delta = [-\phi_\pi, \phi_\pi, -1]'$ and $r_\delta = [1, 0, 0]'$ account for the linear restrictions on δ_t. The conditional likelihood implied by this equation is $\pi_P\,|\tau,\theta_{-\delta} \sim N(X_\delta(R_\delta\delta + r_\delta), \sigma_P^2))$. The posterior is obtained by mul-

tiplying this density with the beta prior distribution. The resulting conditional posterior has bounded support. To generate a draw of δ from this distribution, the density is first evaluated on a fine grid for values of δ. A draw is then generated with the inverse-transform method.

5. Sampling the shock variances

Since the inverse Gamma priors for the variances are conjugate, it is straight forward to sample them from their full conditional distributions. Conditional on the τ and all parameters except for the variances, the errors are obtained via the model relations. For simplicity denote all shocks $\varepsilon = [\varepsilon^g, \varepsilon^\pi, \varepsilon^{MP}, \varepsilon^P, \varepsilon^\star]$. Then, the ith for $i = g, \pi, MP, P, \star$ shock in regime $m = 1, .., M$ is denoted $\varepsilon_{i,m}$. The corresponding prior for each $\sigma_{i,m}^2$ is $\sigma_{i,m}^2 \sim IG(\nu_i, S_i)$. Note that the priors are not regime specific. The full conditional posterior is obtained from standard conjugate results:

$$\sigma_{i,m}^2 | \tau, \, \theta_{-\sigma_{i,m}^2} \sim IG\left(\frac{T_m}{2} + \nu_i, S_i + 0.5 \varepsilon_{i,m}' \varepsilon_{i,m}\right)$$

where T_m is the number of observations in regime m. All variances are sampled individually from their posteriors consecutively.

6. Sampling the initial values τ_0

The initial values for the states τ_0 appear in two places in the model: In the observation equation and in the law of motion for the first τ at $t = 1$, denoted τ_1. Therefore, the full conditional posterior is given by:

$$p(\tau_0 | \, y, \tau, \theta_{-\tau_0}) \propto p(y| \, \tau, \theta) p(\tau_1 \, |\theta) p(\tau_0)$$

where $\theta_{-\tau_0}$ collects all model parameters except the initial states $\tau_{[0:1-q]}$. Note that in the density $p(y| \, \tau, \theta)$ up to q pre-sample values for τ will appear through via the Taylor rule. For sampling τ_0, I make a simplifying assumption that all values in further pre-sample periods are equal to the value at $t = 0$, i.e. $\tau_{-s} = \tau_0$ for $s \geq 1$. Derivation of $p(y| \, \tau, \theta)$ follows from the observation equation where c was inserted

$$y = \gamma_y + \gamma_\tau + H_{A,\beta}^{-1} \gamma_c + (\Gamma + H_{A,\beta}^{-1} \Lambda) \tau + H_{A,\beta}^{-1} e. \tag{2.46}$$

Redefine the following expressions

$$
\gamma_c = \gamma_{c,\tau_0}\tau_0 = \left[\left(\begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_{q-1} & \lambda_q \\ \lambda_2 & \lambda_3 & \dots & \lambda_q & 0 \\ \vdots & & & & \vdots \\ \lambda_q & 0 & \dots & & 0 \end{bmatrix} \begin{bmatrix} \mathbf{1}_{r\times r} \\ \mathbf{1}_{r\times r} \\ \vdots \\ \mathbf{1}_{r\times r} \end{bmatrix}\right)' \quad 0_{r\times(T-q)n} \right]' \tau_0 \tag{2.47}
$$

$$
\gamma_\tau = \gamma_{\tau,\tau_0}\tau_0 = [\Gamma_1', 0_{r\times n(T-1)}]'\tau_0 \tag{2.48}
$$

$$
\alpha_\tau = [F', 0_{r\times n(T-1)}]'\tau_0 \tag{2.49}
$$

and factoring out τ_0 gives

$$
y = \gamma_y + \left(\gamma_{\tau,\tau_0} + H_{A,\beta}^{-1}\gamma_{c,\tau_0}\right)\tau_0 + \left(\Gamma + H_{A,\beta}^{-1}\Lambda\right)\tau + H_{A,\beta}^{-1}e.
$$

Using this reformulation in the conditional distribution in Equation (**??**) implies following conditional likelihood

$$
y \mid \tau_0, \tau, \theta \sim N(m_{y,\tau_0} + X_{y0}\tau_0, \ K_{y|\tau}^{-1})
$$

$$
\text{with} \qquad m_{y,\tau_0} = \gamma_y + (\Gamma + H_{A,\beta}^{-1}\Lambda + H_{A,\beta}^{-1}B)\tau = \gamma_y + X_{y\tau}\tau \tag{2.50}
$$

$$
\text{and} \qquad X_{y,\tau_0} = \gamma_{\tau,\tau_0} + H_{A,\beta}^{-1}(\gamma_{c,\tau_0} - B[F', 0_{r\times n(T-1)}]')
$$

Additionally, the initial values appear in the of motion for τ_t at $t = 1$:

$$
\tau_1 = F_1\tau_0 + J_1 e_1 + O_1 u_1 \tag{2.51}
$$

which implies the distribution (marginal of y) $\tau_1|\theta \sim N(F_1\tau_0, K_{\tau_1}^{-1})$ with precision matrix $K_{\tau_1} = (J_1\Sigma_1 J)_1' + O_1\Omega_1 O_1')^{-1}$. Combining these two densities with a normal prior $\tau_0 \sim N(a_0, B_0)$ yields the following conditional posterior:

$$
\tau_0 \mid y, \tau, \theta_{-\tau_0} \sim N(\hat{\tau}_0, K_{\tau_0}^{-1})
$$

$$
\text{with} \qquad K_{\tau_0} = B_0^{-1} + F'K_{\tau_1}F + (H_{A,\beta}X_{y0})'P_\tau^{-1}H_{A,\beta}X_{y0}
$$

$$
\begin{aligned}
\text{and} \qquad \hat{\tau}_0 &= P_{\tau_0}^{-1}\left(B_0^{-1}a_0 + F'K_{\tau_1}\tau_1 + (H_{A,\beta}X_{y0})'P^{-1}H_{A,\beta}(y - m_{y,\tau_0})\right) \\
&= P_{\tau_0}^{-1}\left(B_0^{-1}a_0 + F'K_{\tau_1}\tau_1 + (H_{A,\beta}X_{y0})'P^{-1}H_{A,\beta}(y - \gamma_y - (\Gamma + H_{A,\beta}^{-1}\Lambda + H_{A,\beta}^{-1}B)\tau)\right) \\
&= P_{\tau_0}^{-1}\left(B_0^{-1}a_0 + F'K_{\tau_1}\tau_1 + (H_{A,\beta}X_{y0})'P^{-1}\left[H_{A,\beta}(y - \gamma_y - \Gamma\tau) - (\Lambda + B)\tau\right]\right)
\end{aligned} \tag{2.52}
$$

Chapter 3

Structural Vector Autoregressions with Common Factor Stochastic Volatility

3.1 Introduction

A large and fast growing literature uses stochastic volatility (SV) to model the time variation in the variances of shocks in structural and reduced-form vector autoregressions (SVAR). The vast majority of this literature assumes that the time variation in each of the variances is driven by its own *independent* stochastic volatility (ISV) process. However, these assumptions are rarely checked. For example, Clark and Ravazzolo (2015) examine the forecast accuracy of a battery of different independent time-varying volatility specifications but fail to consider possible common sources of the time variation in the volatilities.

Yet, the time variation in the variances of various shocks may not be independent from one another. For example, in larger dimensions, it can be expected that changes in volatilities of variables that capture the same economic concept, such as prices or economic activity, are driven by a common volatility factor, see e.g. Carriero et al. (2019). Moreover, ignoring possible common drivers can be problematic when the structure of the volatilities is at the heart of the research question, as in Clark and Davig (2011), which decomposes the decline in the volatility of long-term inflation expectations.

Related literature. In the literature, there are currently two common alternatives to model common time variation in volatilities. The first approach assumes that the time

variation in the volatilities of the residuals of an SVAR is driven by only a limited number of common factors. For example, the common stochastic volatility (CSV) model of Carriero et al. (2016) assumes that one single factor scales the entire residual co-variance matrix. Carriero et al. (2018) allow one factor each for the time variation of the macroeconomic and financial volatility, respectively. However, neither paper checks for the most appropriate number of factors and, *a priori*, restrict factor loadings more heavily than necessary from an econometric perspective.

A second strand of the literature builds on factor models with stochastic volatility (FSV), as in Kastner et al. (2017) and Kastner (2019). Rather than a factor structure in the volatilities, this strand assumes a factor structure in the residuals and allows for stochastic volatility in both the factor innovations and deviations from the factors. Such models are applied to the residuals of SVARs by Carriero et al. (2022) and Chan (2022). However, these also assume the stochastic volatilities to evolve independently.

This paper. Advancing on the current literature, I propose a model that allows for determining the number of common factors driving the time variation in volatilities and to test restrictions on the factor loadings easily. To achieve this, the new model uses a multivariate version of the stochastic volatility model by Chan (2018), which is based on the non-centered parametrization of Frühwirth-Schnatter and Wagner (2010). This parametrization has the advantage that it is straightforward to check for the number of common volatility factors and non-zero loadings via Savage-Dickey density ratios (SDDR), thus avoiding the costly computation of marginal data densities for Bayesian model comparison. To determine the number of factors, only the most general candidate model needs to be estimated. Moreover, the proposed model nests the CSV model of Carriero et al. (2016). To distinguish the new model from the existing CSV model and the FSV of Kastner et al. (2017), I refer to the new model as the common factor stochastic volatility (CFSV) model.

In a simple, illustrative simulation exercise I demonstrate that the CFSV model does a good job at recovering the true number of SV factors and loadings. The CFSV model also performs well at recovering the true volatility structure in a more realistic SVAR simulation. Finally, I revisit the application of Clark and Davig (2011) with the CFSV model. Clark and Davig decompose the decline in the volatility of observed US long-term inflation expectations from the Survey of Professional Forecasters (SPF). To that end, they employ a five-dimensional SVAR including standard macro variables as well as short- and long-term inflation expectations, while allowing for ISV. They find that the volatility of long-term inflation expectations is almost exclusively driven by a decline in the volatility of

the own shock of long-term inflation expectations. However, they do not check for common factors in the volatilities of their identified shocks and, thus, it is possible that the shock to long-term expectations shares a common volatility factor with the shock to short-term inflation expectations or actual inflation. This could change the interpretation of their results. Moreover, the current environment of increased inflation volatility through adverse international supply and energy price shocks makes volatility spill-overs from shocks to inflation to inflation expectations more probable and poses a potential risk to the anchoring of inflation expectations that central banks around the world are monitoring closely.[38] The CFSV model is well suited to uncover such volatility spill-overs. Allowing for CFSV instead of ISV in their SVAR, I find that the stochastic volatility structure of the five dimensional SVAR is best described by just two common factors, a finding that would not have been possible with the more restrictive CSV model of Carriero et al. (2016). The first factor loads only into long-term inflation expectations, thereby supporting the original interpretation of Clark and Davig. However, the second volatility factor loads onto both short-term expectations and actual inflation, thus implying that large shocks to actual inflation are accompanied by larger shocks to short-term inflation expectations. To the extent that short-term inflation expectations are a driver of current inflation, as suggested by most modern macro models, conventional SVARs that ignore volatility spill-overs might underestimate the adverse effect of increased inflation volatility.

Outline. Section 2 introduces the new common factor SV model, discusses estimation, and how to check for the number of common factors. Section 3 includes the simulation exercises and shows that the new model successfully recovers the true volatility structure of a simple DGP with fixed mean and an SVAR. Section 4 uses the new model to check for the number of common SV factors and loadings in the application off Clark and Davig (2011). Section 5 concludes.

3.2 The common factor stochastic volatility model

The CFSV model is a multivariate generalization of the univariate non-centered parametrization for stochastic volatility of Chan (2018). To set the stage, I briefly revisit this specification here. Therefore, let $y_t = [y_{1t}, ..., y_{nt}]$ be a $n \times 1$ data vector. It is assumed that each individual

---

[38]ECB Council member Schnabel (2022) referred to this environment as the *Great Volatility* in her speech at the Jackson Hole Economic Symposium, August 27, 2022.

component $i = 1, ..., n$, evolves according to

$$y_{it} = m_{it} + e^{0.5(\tilde{h}_{i0} + \omega_i \tilde{h}_{it})} \varepsilon_{it} \qquad\qquad \varepsilon_{it} \sim N(0,1) \qquad\qquad (3.1)$$

$$\tilde{h}_{it} = \tilde{h}_{it-1} + u_{it} \qquad\qquad u_{it} \sim N(0,1) \qquad\qquad (3.2)$$

where m_{it} is a time-varying mean and could, for example, represent the regressors of a recursively identified SVAR in A-form. The shocks $y_{it} - m_{it}$ have potentially time-varying variances. In particular, the log-variance $h_{it} = \tilde{h}_{i0} + \omega_i \tilde{h}_{it}$ of shock i consists of a constant $\tilde{h}_{i0}$ and a time-varying part $\tilde{h}_{it}$ that enters through the loading ω_i. For $\omega_i \neq 0$ the variance of shock i is time-varying and constant otherwise. Note that ω_i is identified only up to sign. In contrast to a conventional SV model, this allows for checking the restriction $\omega_i = 0$ easily with the SDDR or by inspection of the posterior density as shown. If the posterior of ω_i is unimodal with a peak at zero, this supports $\omega_i = 0$, while a bimodal distribution with only little density mass at zero is evidence against a constant volatility.

Next, consider the model for the entire data vector y_t

$$y_t = m_t + BD_t \varepsilon_t \qquad\qquad \varepsilon_t \sim N(0, I_n) \qquad\qquad (3.3)$$

$$D_t = \text{diag}\left(e^{0.5(\tilde{h}_0 + \Omega_n \tilde{h}_t)}\right) \qquad\qquad\qquad\qquad (3.4)$$

$$\tilde{h}_t = \tilde{h}_{t-1} + u_t \qquad\qquad u_t \sim N(0, I_n) \qquad\qquad (3.5)$$

where D_t is the $n \times n$ matrix of standard deviations and $\Omega_n = \text{diag}(\omega_1, ..., \omega_n)$ collects the scales of the volatility factors $\tilde{h}_t = [\tilde{h}_{1t}, ..., \tilde{h}_{nt}]'$. Note that e^x denotes the entry-wise exponential function. The entries in the diagonal matrix of standard deviations D_t have a time-varying part $\Omega_n \tilde{h}_t$ and a constant part $\tilde{h}_0$. The matrix B allows for correlations of the residuals across equations. In the SVAR literature, B is referred to as the structural impact matrix, while $D_t \varepsilon_t$ and $BD_t \varepsilon_t$, respectively, are the structural and reduced form shocks. The covariance matrix of the structural shocks is given by $D_t D_t' = \text{diag}\left(e^{\tilde{h}_0 + \Omega_n \tilde{h}_t}\right)$ and the covariance matrix of the reduced form shocks is $\Sigma_t = BD_t D_t' B'$.

So far the time variation in the volatilities of the structural shocks is restricted to be independent from each other because Ω_n is assumed to be diagonal.[39] To relax this assumption and to allow for $r \leq n$ common factors in the time variation of the volatilities, I introduce the potentially full $n \times r$ matrix Γ of (sign-)*normalized factor loadings*. Furthermore, I allow for the possibility that the diagonal matrix Ω_r has a reduced dimension of $r < n$.

---

[39]Note that, since the reduced form shocks are linear combinations of the structural shocks, there are also common components in the volatilities of the reduced form shocks that are made of common structural shocks.

Defining the matrix $\underset{n \times r}{\tilde{\Gamma}} := \underset{n \times r}{\Gamma} \underset{r \times r}{\Omega_r}$, which contains the *unnormalized loadings*, the standard deviations of the structural shocks with common factors can be written as

$$D_t = \text{diag}\left(e^{0.5(\tilde{h}_0 + \tilde{\Gamma}\tilde{h}_t)}\right) \tag{3.6}$$

where the vector of common volatility factors $\tilde{h}_t$ is of dimension $r \times 1$. Since the diagonal elements of the scale matrix Ω_r are identified only up to sign, this also holds for the columns of the unnormalized loadings matrix $\tilde{\Gamma}$. This feature will be exploited to check for zero elements in $\tilde{\Gamma}$. Following the literature on identification in factor models, restrictions on $\tilde{\Gamma}$ are required to identify the factors. For simplicity, I impose a lower triangular structure in the upper $r \times r$ block of $\tilde{\Gamma}$.

Restrictions implied by the CSV model. The CFSV model nests the CSV model of Carriero et al. (2016) and, thus, the implied restrictions can be tested. The CSV model assumes that the variation of the residual covariance matrix is driven by a single factor h_{1t} as follows

$$BD_t\varepsilon_t \sim (0, e^{h_{1t}}BB')$$

where $\Sigma = BB'$ is a constant, symmetrical positive definite matrix. The CFSV model generates the CSV model if $D_t = e^{0.5h_{1t}}I_n$. This amounts to setting $r = 1$ and restricting $h_0 = 0$ and $\tilde{\Gamma} = \mathbf{1}_{n \times 1}$. Hence, the CFSV model is a strict generalization of the CSV model and does not rule out the finding that there is only one common volatility factor.[40]

3.2.1 Estimation

In order to estimate a VAR with common factors in the stochastic volatilities in a Gibbs sampler, one must replace the step that draws from independent stochastic volatilities with a step that takes into account the reduced rank and the non-diagonal loadings matrix $\tilde{\Gamma}$. An additional step is needed to sample the non-zero elements in $\tilde{\Gamma}$.

To estimate the SV processes $\tilde{h}_t$, I use the auxiliary-mixture sampler of Kim et al. (1998), implemented with the precision sampler of Chan and Jeliazkov (2009) as in Chan (2018). To understand the intuition of this approach, define $y_t^\star = \log((B^{-1}(y_t - m_t))^2$ and $\varepsilon_t^\star = \log \varepsilon_t^2$

---

[40]A slightly more general version considered in Carriero et al. (2016) allows for idiosyncratic components in the volatilities but maintains the assumption of only one common factor that loads equally into all volatilities.

and rewrite the CFSV model as

$$y_t^\star = \tilde{h}_0 + \underset{n \times r}{\Gamma} \underset{r \times r}{\Omega_r} \tilde{h}_t + \varepsilon_t^\star. \tag{3.7}$$

It follows from $\varepsilon_{it} \sim N(0,1)$ that $\varepsilon_{it}^\star \sim \log \chi^2(1)$. The model is now a linear state space model with non-normal errors. The auxiliary-mixture sampler approximates the $\log \chi^2(1)$ of $\varepsilon_{it}^\star$ with a seven-component mixture of normal distributions. To that end, an additional latent discrete variable, s_{it}, that indicates the mixture component is introduced and estimated. Conditional on s_{it}, the distribution of the errors is normal; the volatility factors and remaining model parameters can be simulated with standard methods.

Normal priors for the unnormalized. Normally distributed errors imply that the natural conjugate priors for the parameters is the normal distribution. Therefore, I use normal priors with the same variances for the non-zero elements of $\tilde{\Gamma}$ in row i and column j

$$\tilde{\gamma}_{ij} \sim N(0, \sigma_{\tilde{\gamma}}^2) \quad \text{for} \quad i = 1, ..., n \quad \text{and} \quad j \leq i \tag{3.8}$$

with $\sigma_{\tilde{\gamma}}^2 = 1$. The normal prior enables results from linear regression to be used to simulate the elements of $\tilde{\Gamma}$. The computational details are included in the Technical Appendix 3.A.

Implied priors for the normalized factor loadings. Since the signs of the columns of $\tilde{\Gamma}$ are not identified but the signs of the loadings might be of interest, it is useful to back out the implied priors on the normalized loadings in Γ. To pin down the normalized loadings, Γ, one element per column needs to be set to unity. Here, I impose a unit-diagonal on Γ. The remaining normalized loadings are related to unnormalized loadings and the scales by

$$\tilde{\gamma}_{ij} = \begin{cases} \omega_j & \text{for} \quad i = j \\ \gamma_{ij}\omega_j & \text{for} \quad i < j \end{cases}$$

Thus, the induced priors for the elements of Ω_r and Γ, ω_j and $\gamma_{ij} = \frac{\tilde{\gamma}_{ij}}{\omega_j}$ are

$$\omega_j \sim N(0, \sigma_{\tilde{\gamma}}^2) \tag{3.9}$$

$$\gamma_{ij} \sim t_1. \tag{3.10}$$

The prior student-t_1 density for γ_{ij} follows from the ratio of two normal densities with equal variances.

3.2.2 Checking hypothesis on the factor loadings with SDDRs

This section shows how to check hypotheses about individual and joint loadings using Savage-Dickey density ratios (SDDR). The use of SDDRs for comparison of nested models avoids the computation of the marginal data density of the two models under comparison. This is particularly advantageous when the computation of the marginal data density is very computationally expensive, as is typically the case with nonlinear state-space models, e.g. stochastic volatility models. Instead of explicitly computing the marginal data densities, the SDDR allows for testing nested models by comparing prior and posterior densities evaluated at parameters values implied by the restriction. For the CFSV model, such restrictions also include the omission of volatility factors. Therefore, only the most general model has to be estimated to determine the number of factors. If the jth component of $\tilde{h}_t$ can be omitted this implies a zero-column at position j in $\tilde{\Gamma}$ and there are only $r-1$ common factors driving the time variation of the volatilities. Assuming a lower-triangular structure for $\tilde{\Gamma}$, checking whether the jth factor can be omitted corresponds to the hypothesis

$$H_0^j : \tilde{\gamma}_{jj} = ... = \tilde{\gamma}_{nj} = 0 \tag{3.11}$$

Checking this hypothesis for all $j = 1, ..., r$ allows for pinning down the number of factors. In addition to the number of factors, it is also of interest which shocks share a common volatility and which do not. This corresponds to checking whether factor j does not load into the volatility of variable i and leads to the following hypothesis on a single loading:

$$H_0^{ij} : \tilde{\gamma}_{ij} = 0 \quad \text{for} \quad i \leq j \tag{3.12}$$

Assessing hypotheses in a Bayesian framework is done by computing the ratio of the marginal likelihoods for the unrestricted and models M_{UR} and M_R, the so-called Bayes Factor:

$$BF_{UR} = \frac{p(y|M_{UR})}{p(y|M_R)} \tag{3.13}$$

The subscript UR for the Bayes Factor indicates that larger values of BF_{UR} indicate evidence in favor of the unrestricted model. Computation of the marginal data density for stochastic volatility models is computationally costly because it requires to numerically integrate out the stochastic volatilities, see e.g. Chan and Eisenstat (2018). However, the parametrization of the model allows for using the SDDR as a short cut to estimating Bayes Factors for the two hypotheses above. Using the SDDR is admissible if the priors for the

parameters involved in the restriction are independent from the remaining parameters.[41]
The independent normal priors for all $\tilde{\gamma}_{ij}$'s satisfy this condition.

The SDDR is the ratio of the prior and the posterior density evaluated at H_0. Using the
SDDR for hypothesis H_0^j of a zero column at position j the Bayes Factor can be computed
as

$$BF_{UR}^{j} = \frac{p(\tilde{\gamma}_{jj} = ... = \tilde{\gamma}_{nj} = 0)}{p(\tilde{\gamma}_{jj} = ... = \tilde{\gamma}_{nj} = 0|y)} \tag{3.14}$$

and for the individual hypothesis H_0^{ij} as

$$BF_{UR}^{ij} = \frac{p(\tilde{\gamma}_{ij} = 0)}{p(\tilde{\gamma}_{ij} = 0|y)} \tag{3.15}$$

where $p(\,\cdot\,)$ and $p(\,\cdot\,|\,y)$ denote prior and posterior densities, respectively. The above ratio
of densities shows that if BF_{UR} becomes larger than one, then the restriction is less likely
to hold under the posterior than under the prior. This is interpreted as evidence against
the restriction and in favor of the unrestricted model because the restriction is less likely
when the data y are taken into account, i.e. under the posterior. In practice the log Bayes
Factor is often reported. A positive $logBF_{UR}$ is evidence against the restriction. Kass and
Raftery (1995) provide guidance as to what magnitudes of Bayes factors should be viewed
as weak or strong evidence.

The individual and joint posterior densities required to compute the SDDR are relatively
straight forward to evaluate from the output of a Gibbs sampler; see Chan (2018). Appendix
3.B contains the details for the computation of the posterior densities for the SDDR for the
CFSV model. The prior densities can be evaluated directly.

Furthermore, Chan (2018) notes that the exact value of the Bayes Factor might be not
very precisely estimated by the SDDR if the denominator, i.e. the posterior density, is close
to zero. Therefore, the Bayes Factor should be complemented by inspecting the shape of
the posterior of each $\tilde{\gamma}_{ij}$ for $i = j, ..., n$. A unimodal posterior of $\tilde{\gamma}_{ij}$ is evidence in favor H_0
and a bimodal posterior is evidence against H_0.

---

[41]Independent priors are a sufficient condition; see the proof in Verdinelli and Wasserman (1995).

3.3 Simulation evidence

3.3.1 A simple simulation exercise

To show that the model can correctly estimate the loadings and common SV processes, I first conduct a simulation exercise based on a simple data generating process (DGP) with fixed mean m, $n = 2$ variables , and $T = 200$ observations. In particular, I use the following matrices:

$$\Omega = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} 1 & 0 \\ \gamma_{21} & 1 \end{pmatrix}, \quad m_t = m = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, h_0 = \begin{pmatrix} -0.1 \\ -0.3 \end{pmatrix}.$$

Furthermore, I set $\omega_1 = 0.15$, $\omega_2 = 0$, and $\gamma_{21} = 1$. Due to $\omega_2 = 0$, there is just one common volatility factor that loads into both variables through $\gamma_{21} \neq 0$. The unnormalized loadings matrix is given by $\tilde{\Gamma} = \begin{pmatrix} \tilde{\gamma}_{11} & 0 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} \end{pmatrix} = \begin{pmatrix} 0.15 & 0 \\ 0.15 & 0 \end{pmatrix}$.

I estimate two models on the same simulated data from this DGP: the CFSV and a conventional model that assumes independent SV processes. Independent SV imposes the restriction $\tilde{\gamma}_{21=0}$. Table 3.3.1 shows the log Bayes Factors from the two models in favor of individual unrestricted loadings $\tilde{\Gamma}$, i.e. against the zero restrictions on the factor loadings. Following the suggestion of Kass and Raftery (1995), log Bayes Factors that exceed 5 are interpreted as decisive evidence against the restriction. Negative values indicate evidence in favor of the restriction. The log Bayes Factors obtained from the CFSV model correctly indicate that $\tilde{\gamma}_{21}, \tilde{\gamma}_{11} \neq 0$, and $\tilde{\gamma}_{22} = 0$. In contrast, applying the ISV model that does not allow for $\tilde{\gamma}_{21} \neq 0$ would support the false conclusion that $\tilde{\gamma}_{22} \neq 0$ because it yields a large Bayes Factor against the restriction $\tilde{\gamma}_{22} = 0$.

Table 3.3.1: Checking restrictions on the factor loadings in the simple simulation

Model	Restriction under H_0^{ij}		
	$\tilde{\gamma}_{11} = 0$	$\tilde{\gamma}_{21} = 0$	$\tilde{\gamma}_{22} = 0$
CFSV	37.06	10.32	-1.99
ISV	26.44	–	33.81

Notes: Reported figures are log-Bayes Factors in favor of the unrestricted model, i.e. values larger (smaller) than zero indicate evidence against (in favor of) the restriction. CFSV denotes log Bayes Factors from the new common factor SV model and ISV denotes the conventional independent SV model.

Following the suggestion of Chan (2018), Figure 3.3.1 compares the posterior draws from both models with the prior densities and the true values. It confirms the conclusions,

Figure 3.3.1: Posterior densities of the unnormalized loadings in the CFSV and ISV models: Simple DGP
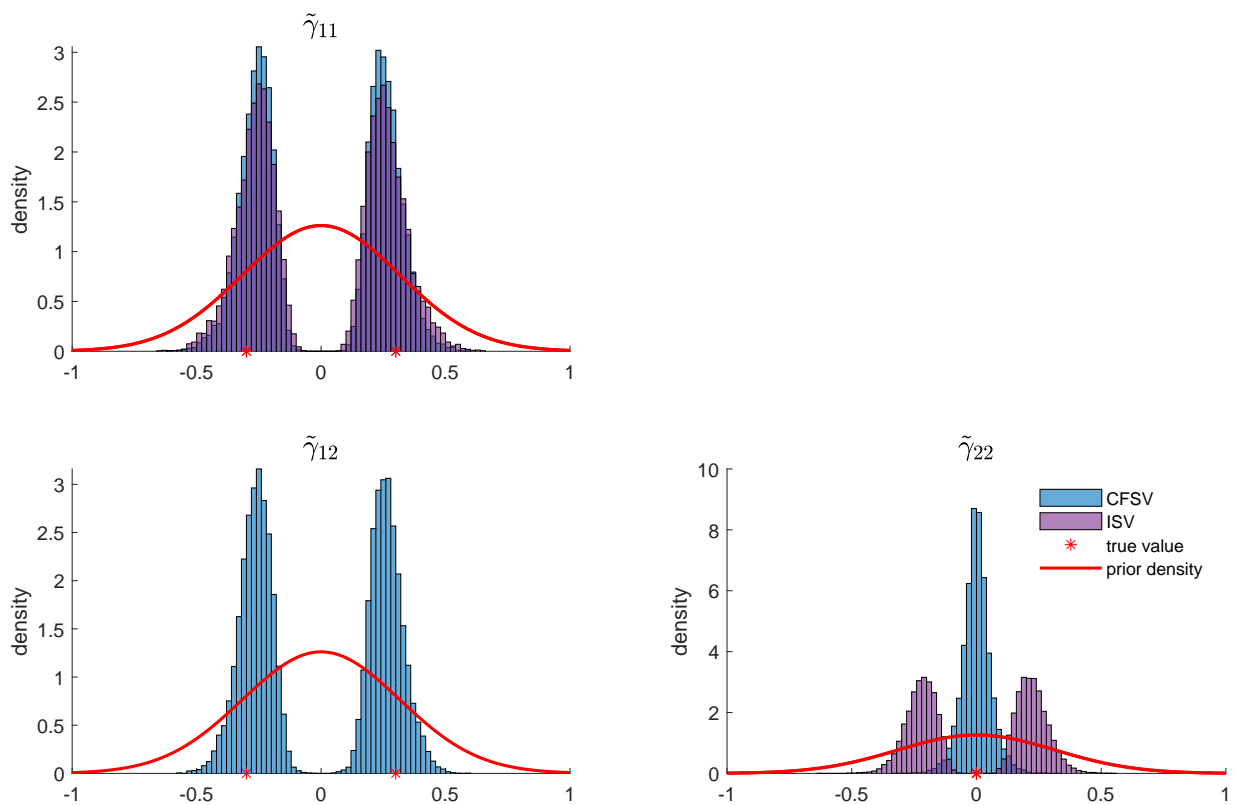
Notes: Gaussian prior densities and histograms of the posterior draws for unnormalized the SV loadings $\tilde{\Gamma}$ from estimation of the CFSV and ISV model on the same simulated data from the simple DGP.

reached by the log Bayes Factors. Both models deliver bimodal posterior distributions for $\tilde{\gamma}_{11}$, indicating $\tilde{\gamma}_{11} \neq 0$. Moreover, the histograms from both models lie virtually on top of each other. For $\tilde{\gamma}_{22}$, the posterior from the CFSV model has a single mode at zero, indicating $\tilde{\gamma}_{22} = 0$. In contrast, the ISV model delivers a bimodal posterior supporting the false conclusion of a non-zero $\tilde{\gamma}_{22}$ from the log Bayes Factors.

The next section examines how the CFSV model performs in a simulation exercise with a SVAR that is more relevant for applied research than the very simple exercise of this section.

3.3.2 SVAR Simulation

To examine how the new volatility model performs with a model that is routinely used in empirical research, this section conducts a simulation with an SVAR DGP. The recursive SVAR for the $n = 3$ vector y_t and volatility factors $\tilde{h}_t$ reads

$$A_0 y_t = A_1 y_{t-1} + e^{0.5(\tilde{h}_0 + \tilde{\Gamma}\tilde{h}_t)} \varepsilon_t, \qquad \varepsilon_t \sim N(0, I_n)$$
$$\tilde{h}_t = 0.95 \tilde{h}_{t-1} + u_t, \qquad u_t \sim N(0, I_n)$$

with coefficient matrices

$$A_0 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \frac{4}{3} & -1 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0.79 & 0 & 0.25 \\ -1.39 & 0.95 & -0.96 \\ 0.98 & -0.95 & 1.41 \end{bmatrix}, \quad \tilde{\Gamma} = \begin{bmatrix} \tilde{\gamma}_{11} & 0 & 0 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} & 0 \\ \tilde{\gamma}_{31} & \tilde{\gamma}_{32} & \tilde{\gamma}_{33} \end{bmatrix}, \quad h_0 = \begin{bmatrix} -1 \\ -1.5 \\ -1.25 \end{bmatrix}$$

and $\tilde{\gamma}_{11} = \tilde{\gamma}_{21} = 0.3$, $\tilde{\gamma}_{22} = \tilde{\gamma}_{32} = 0.5$, $\tilde{\gamma}_{31} = \tilde{\gamma}_{33} = 0$. Since $\tilde{\gamma}_{33} = 0$, the third volatility factor does not enter the volatility of any shock. As a result, the column rank of $\tilde{\Gamma}$ is reduced to $r = 2$. Therefore, the simulation shows whether the estimation can reliably recover a reduced rank in a more complex model. The volatility processes $\tilde{h}_t$ follow stationary AR(1) processes in order to ensure a constant unconditional variance matrix of the shocks and stable simulated paths for $\tilde{h}_t$.

The recursively identified SVAR with CFSV is estimated by replacing the steps to sample the independent SV in the algorithm of Chan and Eisenstat (2018) with the steps for the CFSV model outlined in Appendix 3.A.[42] The prior distribution for the VAR parameters $b = [vec(A_1)' \ vech(A_0)']'$ is set to a relatively uninformative normal distribution $b \sim N(0, I)$.

---

[42]To implement the Gibbs steps for the CFSV model, I first obtain the MATLAB function `SVRW.m` that implements one Gibbs draw from an individual SV process with random walk law of motion from the homepage of Joshua Chan and modify it accordingly.

In general, the prior for b will not dramatically influence the ability of the model to recover the factor structure in the volatilities. The priors for volatility loadings are more important. Therefore, I choose a prior that is not too far off from the true values $\tilde{\gamma}_{ij} \sim N(0, 0.2)$, and a very wide prior for the initial values $\tilde{h}_0 \sim N(0, 10)$. As emphasized by Chan (2018) and shown in Kroese et al. (2014), the implied prior for $\tilde{\gamma}_{ij}^2$ is an inverse gamma distribution with parameters $IG(\frac{1}{2}, \frac{1}{2V_\gamma})$ and mean V_γ. Therefore, the normal prior is reasonably close to the true parameters. To make the simulation comparable to real life research situation, I do not use the true values for specifying the prior. To examine how well the model CFSV model recovers the true number of volatility factors, I simulate 500 data sets from the DGP and estimate the model on each of them.

The estimation recovers the correct number of volatility factors and also the magnitude of the loadings. Figure 3.3.2 shows the distribution of the estimated log Bayes Factors in favor of $\tilde{\gamma}_{ij} \neq 0$. Recall that values larger than zero indicate evidence in favor of $\tilde{\gamma}_{ij} \neq 0$. The log Bayes Factors for $\tilde{\gamma}_{11}$, $\tilde{\gamma}_{22}$ and $\tilde{\gamma}_{32}$ always have the correct, positive sign. For $\tilde{\gamma}_{21}$ the log Bayes Factors are negative only in a small number of cases, indicating that this is the most challenging loading to recover. For the zero elements $\tilde{\gamma}_{31}$ and $\tilde{\gamma}_{33}$, the majority of values is negative, and if positive only slightly so. In general, these results show that the model is successful in delivering the correct number of volatility factors $r = 2$. The next section applies this analysis to US macro data.

Figure 3.3.2: Distribution of log Bayes Factors in favor $\gamma_{ij} \neq 0$ for the elements of $\tilde{\Gamma}$ from the VAR simulation
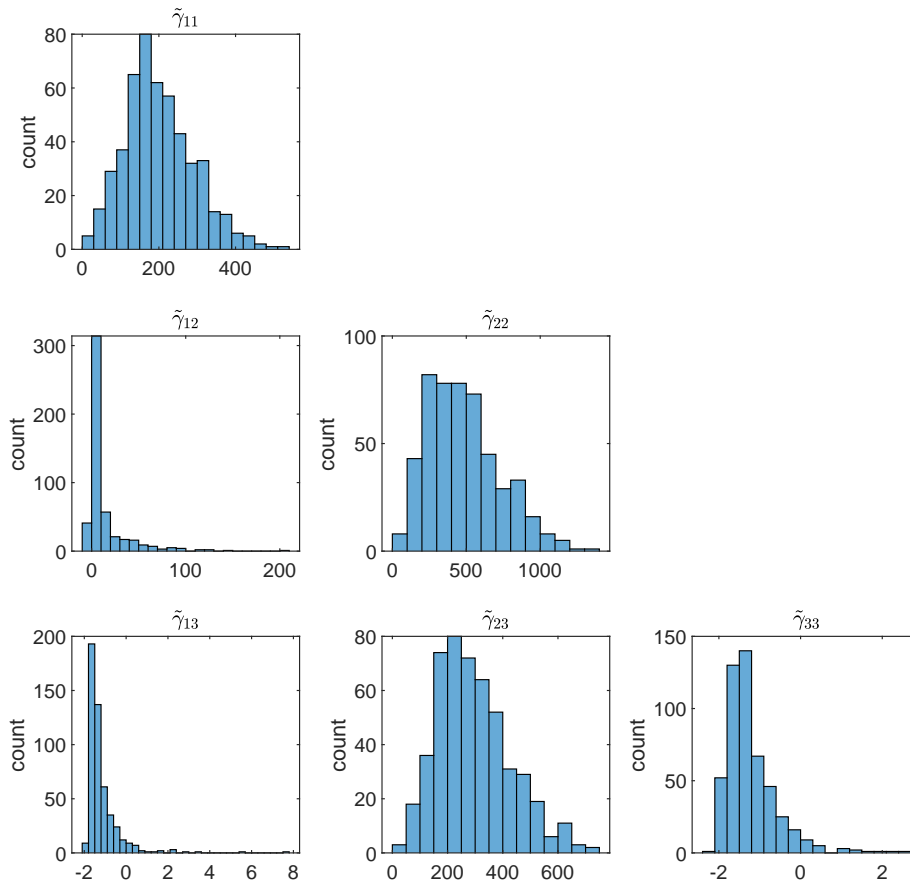
3.4 Common factors in the volatility of US inflation expectations

In this section, I use the CFSV model to revisit the application of Clark and Davig (2011), henceforth CK2011. The aim of their application is to decompose the time variation in the volatility of US long-term inflation expectations. Using an SVAR with ISV, they find that the stochastic volatility of the own shock of long-term inflation expectations ε_t^{LT} explains the bulk of the time variation in the volatility of long-term inflation expectations. However, due to the *a priori* assumption of independently evolving SV, they cannot *check* whether the volatility of ε_t^{LT} is also driven by factors that are common to the volatilities of the other shocks. If the volatility of ε_t^{LT} were indeed driven by common factors, the interpretation of Clark and Davig might not hold. Moreover, the current environment of increased inflation volatility through adverse effects from international supply chain disruptions following the COVID-19 pandemic and energy price shocks from the war in Ukraine makes volatility spill-overs to inflation expectations more probable. This poses a potential risk to the anchoring of inflation expectations that central banks around the world are monitoring closely. The CFSV model is well suited to uncover such volatility spill-overs.

SVAR specification. Clark and Davig estimate a Bayesian SVAR with $p = 4$ lags, with time-varying parameters and individual stochastic volatility on 108 quarterly observations from 1981Q3 to 2008Q2. The vector of the $n = 5$ endogenous variables is

$$y_t = \begin{bmatrix} \pi_t^{LT} & \pi_t^{ST} & \pi_t & cfnai_t & ffr_t \end{bmatrix}'.$$

Long- and short-term inflation expectations, denoted π_t^{LT} and π_t^{ST}, are measured by 9 year 1 year ahead, and 1 year ahead inflation forecasts from the Survey of Professional Forecasters (SPF) by the Philadelphia Fed.[43] $cfnai_t$ denotes the Chicago Fed National Activity Index and is a proxy for the output gap. It is the first factor obtained from a large number of macroeconomic time series and available from FRED. ffr_t is the federal funds rate. Clark and Davig recursively identify five shocks $\varepsilon_t = [\varepsilon_t^{LT}, \varepsilon_t^{ST}, \varepsilon_t^{\pi}, \varepsilon_t^{cfnai}, \varepsilon_t^{ffr}]$

---

[43]Only 1 year and 10 year ahead expectations are directly available from the SPF. To account for the overlap of the 1 year ahead expectation that is also contained in the original 10 year SPF expectation measure, I follow Clark and Davig and compute:

$$9 \text{ year 1 year ahead SPF expectation} = \frac{10 \text{ year SPF expectation} - 0.1 \times 1 \text{ year SPF expectation}}{0.9}.$$

and allow for independent stochastic volatility for each shock.

I re-estimate a constant parameter version of their model with CFSV as in (3.7) instead of ISV.[44] For the sake of comparability, I follow CK2011 exactly in all other modeling aspects. To identify the factors, I impose a lower triangular structure on the unnormalized factor loadings $\tilde{\Gamma}$, as in the preceding sections. This implies that only the first SV factor loads onto ε_t^{LT}. Hence, if the first factor also loads onto the volatility of any other shock, i.e. $\tilde{\gamma}_{i1} \neq 0$ for $i = 2, ..., n$, this would cast doubt on the interpretation of Clark and Davig. By contrast, a finding that the first factor loads exclusively onto the volatility of ε_t^{LT} and not onto the volatilities of any other shock confirms the original interpretation. In this case, the volatility of ε_t^{LT} is entirely driven by its own independent factor, as in the original model.

Determining the number of factors. The model is estimated allowing for the maximum of $r = 5$ SV factors. To determine the number of factors, the exclusion of the jth factor is tested for all $j = 1, ..., r$ with the joint hypothesis in (3.11). Table 1 shows the resulting log Bayes Factors from the joint hypothesis in favor of the unrestricted models. The last row shows that under $r = 5$, only the first two factors are required, because the log Bayes Factors are positive only for $j = 1, 2$. To ensure that this finding holds when allowing for fewer factors, I re-estimate the model for $r = 1, ..., 4$ and compute the log Bayes factors for the joint hypothesis. The other rows in Table 1 confirm that only the first two SV factors are required to model the time variation of the five shocks, for all possible values of r.

Assessing the individual factor loadings. Figure 3.4.1 shows the prior and posterior densities of the individual elements of $\tilde{\Gamma}$ from the most general model with $r = 5$ along with the individual log Bayes Factors $\log BF_{UR}$ in favor of the unrestricted model, i.e. against $H_0^{ij} : \tilde{\gamma}_{ji} \neq 0$. The individual log Bayes Factors confirm that only the first two SV factors are required, as positive $\log BF_{UR}$ only occur for $j = 1, 2$; see the first two columns of Figure 3.4.1. This is also confirmed by inspecting the posterior densities. Only for $\tilde{\gamma}_{11}$, $\tilde{\gamma}_{22}$, and $\tilde{\gamma}_{23}$ are the posterior densities clearly bimodal and lower than the prior density at zero. For $\tilde{\gamma}_{25}$ the result is less clear: Despite a bimodal posterior, there remains substantial posterior density mass at zero. Most relevant for the application of CK2011, the first factor does not load into any of the other shocks, because the posteriors densities of $\tilde{\gamma}_{i1}$ for $i = 2, ..., 5$ peak at zero. This confirms the interpretation of Clark and Davig (2011) that the volatility of shocks to long-term inflation expectations is mainly driven by its *own* volatility factor.

---

[44]Clark and Davig show thar their results hold also in a constant parameter version of their model, so I do not allow for time variation in the VAR parameters as they are not at the heart of the analysis.

Table 3.4.1: Exclusion of the jth SV factor for various r

| | $H_0^j : \tilde{\gamma}_{jj} = ... = \tilde{\gamma}_{nj} = 0$ for $j = 1, ..., r$ | | | | |
	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
r=1	15.02				
r=2	58.48	11.79			
r=3	56.93	9.31	-1.30		
r=4	25.39	0.30	-1.82	-2.02	
r=5	34.38	7.03	-1.52	-1.05	-0.43

Notes: Reported figures are log Bayes Factors in favor of the unrestricted model, i.e. positive values indicate evidence for the unrestricted model and against H_0^j. r is number of factors in each model. j is the factor that is tested under H_0^j. The strength of evidence for log Bayes Factor according to Kass and Raftery (1995) is: $0 < \log BF < 1$: not worth mentioning, $1 < \log BF < 3$: positive, $3 < \log BF < 5$: strong, $5 < \log BF$: very strong.

Figure 3.4.1: Unnormalized factor loadings $\tilde{\Gamma}$ of the CFSV factors for $r = 5$ in the CK2011 Application
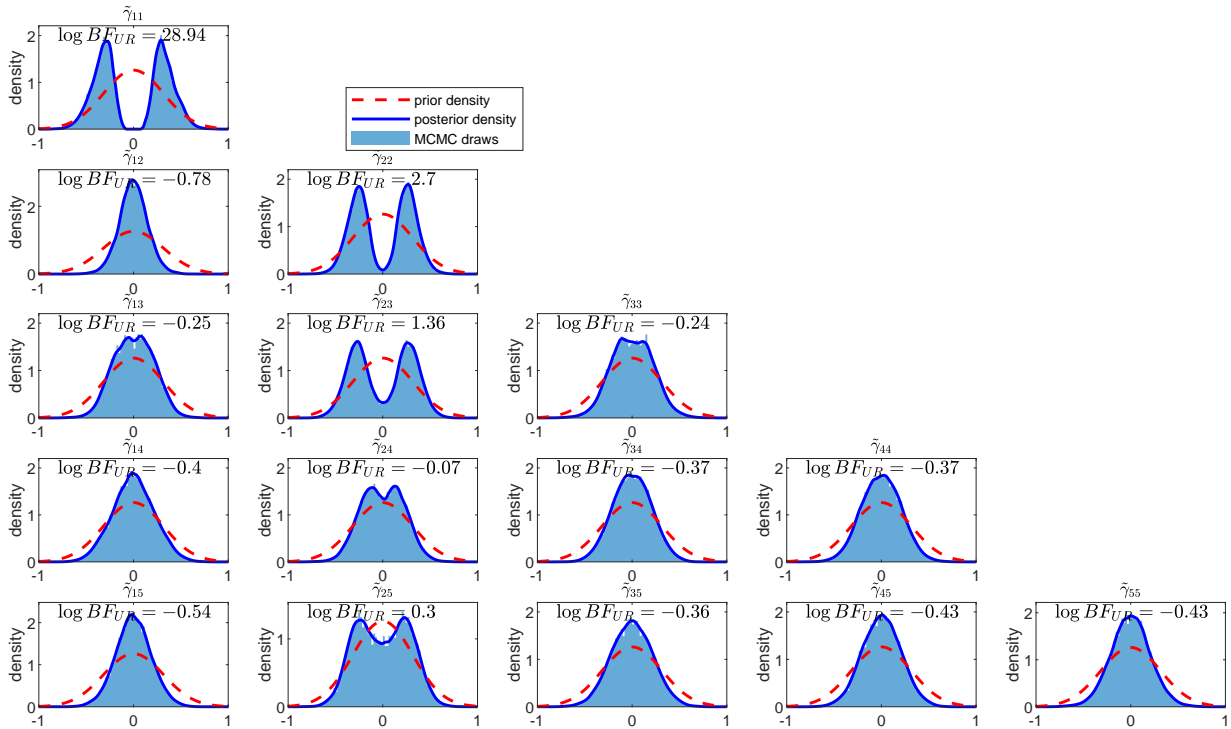
Notes: Gaussian prior densities and simulated posterior densities of the elements in $\tilde{\Gamma}$ with the individual log Bayes Factor BF_{UR} in favor of the unrestricted element $\tilde{\gamma}_{ji} \neq 0$.

However, Figure 3.4.1 also reveals that the second factor simultaneously loads onto the volatilities of the shocks to short-term inflation expectations and actual inflation through the $\tilde{\gamma}_{22}$ and $\tilde{\gamma}_{32}$. Although the log Bayes Factor in favor of $\tilde{\gamma}_{32} \neq 0$ does not indicate very strong evidence, the possibility of such volatility spill-overs may be important for assessing the risk of a de-anchoring of short-term inflation expectations. The posterior distribution of the normalized loading of the second factor to actual inflation volatility, computed as $\gamma_{32} = \frac{\tilde{\gamma}_{32}}{\tilde{\gamma}_{22}}$ has a mean of 0.99 and the 5% and 95% quantiles at 0.19 and 2.65.[45] This indicates that an increase in the second factor on average suggests an equal increase in the volatility of shocks to short-term inflation expectations and inflation. At the current juncture, with increased inflation volatility, this implies that, first, the range of possible outcomes for short-term inflation expectations may be larger than otherwise expected due to larger own shocks to short-term inflation expectations; i.e. expectations might be less well anchored. Second, to the extent that actual inflation is driven by short-term inflation expectations as in the conventional New Keynesian Phillips Curve framework, the higher variability in inflation expectations can add to movements of actual inflation.

3.5 Conclusion

This chapter proposes a new model for stochastic volatility that allows a flexible number of common factors to drive the time variation in the volatilities of multiple shocks. The model relies on a multivariate non-centered parametrization of the SV processes with a triangular factor structure. The number of factors and the loadings can be checked using Savage-Dickey density ratios. Simulation evidence suggests that the model does a good job in recovering the true number of factors. Using the new model, I revisit the application of Clark and Davig (2011), who find that, in a SVAR with independent SV, the time variation in the volatility of US long-term inflation expectations is predominantly driven by the volatility of its own shock. The CFSV model prefers a reduced number of only two factors that drive the volatilities of the five shocks in the SVAR. Nonetheless, the CFSV model confirms that the volatility of the shock to long-term inflation expectations is exclusively driven by its own volatility factor. However, the second factor drives both the volatility of shocks to actual inflation and short-term inflation expectations. At the current juncture, with increased inflation volatility, this implies potentially an additional source of inflation variability that is overlooked by models with independent stochastic volatility.

---

[45]In contrast, for all other normalized loadings the inter-quantile range of the 5% and 95% quantiles covers zero.

One caveat of the new model is that the finding of common factors in volatilities depends crucially on the identification of the shocks. In the recursively identified SVAR of Clark and Davig, this implies dependence on the ordering of the variables. Bertsche and Braun (2022) and Chan et al. (2022) propose algorithms for the identification of SVARs via stochastic volatility that is independent of the variable ordering. Yet, testing for the number of volatility processes that are required for identification is still an active area of research. Thus, a combination of the CFSV model proposed in this paper with identification via stochastic volatility could help solving both issues and it would be interesting to explore this in future research.

Appendix

3.A   Gibbs sampler for the CFSV model

Sampling the SV factors h. Since the model allows the volatility factors $\tilde{h}$ to appear in more than just one equation, they are sampled jointly. Otherwise the procedure is exactly equal to the auxiliary mixture sampler of Kim et al. (1998), as applied in Chan (2018).

First, it follows from $\varepsilon_{it} \sim N(0, 1)$ that $log(\varepsilon_{it}^2)$ has a $\log \chi^2(1)$ distribution. Kim et al. (1998) show that the $\log \chi^2(1)$ distribution can be well approximated using a seven-component Gaussian mixture density with fixed parameters. By introducing an indicator variable for the mixture component indicators, $s_{it} = \{1, ..., 7\}$ for $t = 1, ..., T$ and $i = 1, ..., n$, the errors of (3.7) can be approximated by a conditionally Gaussian distribution $\varepsilon_{it}^\star \sim N(\mu_{s_i t}, \sigma_{s_i t}^2$ where $\mu_{s_i t}$ and $\sigma_{s_i t}^2$ are the parameters of the corresponding component of the Gaussian mixture approximation to the $\log \chi^2(1)$ distribution; see Table 4 in Kim et al. (1998). Conditional on the mixture indicator $s = [s_1', ..., s_T']'$ with $s_t = [s_{1t}, ..., s_{nt}]'$ for all t, the state space model in (3.7) is conditionally linear and Gaussian, thus sampling h can be done efficiently with the precision sampler of Chan and Jeliazkov (2009).

To that end, stack up (3.7) over T to get the model for $y^\star = [y_1^{\star\prime}, ..., y_T^{\star\prime}]'$ and $h = [h_1', ..., h_T']'$ as

$$ y^\star = \tilde{h}_0 \otimes \mathbf{1}_T + (\tilde{\Gamma} \otimes I_T)\tilde{h} + \varepsilon^\star, \quad \varepsilon^\star \sim N(\mu_s, \Sigma_s) \tag{3.16} $$

with $\mu_s = [\mu_{s_1} n, ..., \mu_{s_T} n]'$ and $\mu_{s_t} = [\mu_{s_1 t}, ..., \mu_{s_n t}]'$. The diagonal variance matrix $\Sigma_s = \mathrm{diag}(\sigma_s^2)$ is constructed accordingly with $\sigma_s^2 = [\sigma_{s_1}^2{}', ..., \sigma_{s_T}^2{}']'$ and $\sigma_{s_t}^2 = [\sigma_{s_1 t}^2, ..., \sigma_{s_n t}^2]'$. μ_s and σ_s^2 are both of dimension $Tn \times 1$. Thus, the conditional likelihood of $y^\star$ is given by

$$ (y^\star | s, \tilde{h}_0, \tilde{h}, \tilde{\Gamma}) \sim N(\mu_s + \tilde{h}_0 \otimes \mathbf{1}_T + (\tilde{\Gamma} \otimes I_T)\tilde{h}, \Sigma_s) \tag{3.17} $$

The prior from $\tilde{h}$ is implied from the random-walk law of motion for $\tilde{h}$ stacked up over T

$$H\tilde{h} = u, \quad u \sim N(0, I_{rT})$$

It follows that $\tilde{h} \sim N(0, (H'H)^{-1})$. Combining the conditional likelihood with the prior yields the conditional posterior

$$(\tilde{h} \,|\, y^\star s, \tilde{h}_0, \tilde{\Gamma}) \sim N(\hat{\tilde{h}}, K_{\tilde{h}}^{-1}). \tag{3.18}$$

with $K_{\tilde{h}}^{-1} = H'H + (\tilde{\Gamma} \otimes I_T)'\Sigma_s^{-1}(\tilde{\Gamma} \otimes I_T)$ and $\hat{\tilde{h}} = K_{\tilde{h}}^{-1}\left((\tilde{\Gamma} \otimes I_T)'\Sigma_s^{-1}(y^\star - \mu_s - \tilde{h}_0 \otimes \mathbf{1}_T)\right)$. It is straight forward to generate a draw from this distribution with the precision sampler of Chan and Jeliazkov (2009).

Sampling the mixture component indicator s. The procedure to sampling the mixture component is not changed compared to the original auxiliary mixture sampler Kim et al. (1998). A textbook style treatment of the routine as applied in Chan (2018) can be found in the teaching notes on Joshua Chans homepage[46].

Sampling the parameters in $\tilde{h}_0$ and $\tilde{\Gamma}$ equation-wise Conditional on the $\tilde{h}$ and all other parameters, (3.7) becomes a set of independent regressions. Assuming independent priors, the elements in $\tilde{h}_0$ and $\tilde{\Gamma}$ can be sampled equation-wise. To see this, rewrite (3.7) in SUR form.

$$y_t^\star = X_t^\star \beta + \varepsilon_t^\star \tag{3.19}$$
$$\beta = vec([\tilde{h}_0, \tilde{\Gamma}]') \tag{3.20}$$
$$X_t^\star = I_n \otimes [1,' \tilde{h}_t'] \tag{3.21}$$

where β obeys the zero restrictions implied by the lower-triangular structure of $\tilde{\Gamma}$. Equation i of the above SUR-system takes the form

$$y_{it}^\star = x_{it}^{\star\,\prime}\tilde{\beta}_i + \varepsilon_{it}^\star \tag{3.22}$$
$$\tilde{\beta}_i = [h_{i0}, \tilde{\gamma}_{i1}, ..., \tilde{\gamma}_{ij}]' \tag{3.23}$$
$$x_{it}^{\star\,\prime} = [1, h_{1t}, ..., h_{1j}] \tag{3.24}$$

---

[46]https://joshuachan.org/notes_BayesMacro.html

where, due to the lower triangular structure in $\tilde{\Gamma}$, the ith element of $y_{it}^{\star}$ is regressed on the first $j = min(r, i)$ elements in $\tilde{h}_t$ and a constant. Stacking up over T yields

$$y_i^{\star} = x_i^{\star}\tilde{\beta}_i + \varepsilon_i^{\star}, \quad \varepsilon_i^{\star} \sim N(\mu_{s_i}, \Sigma_{s_i}) \tag{3.25}$$

with $y_i^{\star} = [y_{i1}^{\star}, ..., y_{iT}^{\star}]'$, $x_i^{\star} = [x_{1t}^{\star}, ..., x_{iT}^{\star}]'$ and $\varepsilon_i^{\star} = [\varepsilon_{i1}^{\star}, ..., \varepsilon_{iT}^{\star}]'$. The error distribution depends on the mixture components s and the indicator for the ith equation and all t, hence, $\mu_{s_i} = [\mu_{s_i1}, ..., \mu_{s_iT}]$ and $\Sigma_{s_i} = \text{diag}(\sigma_{s_i}^2)$ with $\sigma_{s_i}^2 = [\sigma_{s_i1}, ..., \sigma_{s_iT}]$.

With the normal prior $\beta_i \sim N(\beta_{0i}, V_{\beta})$, the conditional posterior for β_i is given by standard linear regression results as

$$\beta_i | \sim N(\hat{\beta}_i, D_{\beta_i}) \tag{3.26}$$

with $D_{\beta_i} = \left(V_{\beta}^{-1} + x_i^{\star\prime}\Sigma_{s_i}^{-1}x_i^{\star}\right)^{-1}$ and $\hat{\beta}_i = D_{\beta_i}\left(V_{\beta}^{-1}\beta_{0i} + x_i^{\star\prime}\Sigma_{s_i}^{-1}(y_i^{\star} - \mu_{s_i})\right)$.

3.B Evaluating the posterior densities for the SDDR

Evaluation of the individual densities $p(\tilde{\gamma}_{ij} = 0 \,|\, y^{\star})$ follows the exposition in Appendix A of Chan (2018) but is repeated here for completeness. From the conditional posterior of β_i in (3.26), it follows that the marginal distribution of $\tilde{\gamma}_{ij}$, given $y^{\star}$ is $N(\hat{\tilde{\gamma}}_{ij}, d_{\beta_i, j+1j+1})$, where $\hat{\tilde{\gamma}}_{ij}$ is the $j + 1$th element of β_i and $d_{\beta_i, j+1j+1}$ is the $j + 1$th element on the diagonal of D_{β_i}. Hence, the marginal density $p(\tilde{\gamma}_{ij} = 0 \,|\, y^{\star}, s, \tilde{h})$ can be evaluated exactly. Note this density is still a conditional density, because it is depending on the volatility factors $\tilde{h}$, the mixture component indicator s, as well as the other model parameters contained in $y^{\star}$. Since these elements are not involved in the restriction, they must be integrated out to obtain $p(\tilde{\gamma}_{ij} = 0 \,|\, y)$, which only depends on the parameter $\tilde{\gamma}_{ij}$ evaluated at the restriction and the data vector y. Integrating out all these elements is done by computing the Monte Carlo average of $p(\tilde{\gamma}_{ij} = 0 \,|\, y^{\star}, s, , h)$ over draws of s and $\tilde{h}$ and the unknown parameters.

To compute the joint density $p(\tilde{\gamma}_{jj} = ... = \tilde{\gamma}_{nj} = 0 \,|\, y^{\star})$ for an entire column of zeros at position $j \leq r$ of $\tilde{\Gamma}$, note that conditional on s and $\tilde{h}$, the individual elements $\tilde{\gamma}_{ij}$ and $\tilde{\gamma}_{kj}$ at the ith and kth row in the jth column of $\tilde{\Gamma}$ are independent from each other for $i \neq k$. This is indicated by the fact that the parameters $\tilde{\Gamma}$ are sampled equation-by-equation, i.e. row-wise. Therefore, the parameters in different rows of $\tilde{\Gamma}$ i.e. the same column are independent from one another. This becomes clear as $\tilde{\gamma}_{kj}$ does not appear in the conditional posterior of an element $\tilde{\gamma}_{ij}$ in the same column j but a different row k. Independence implies

that the conditional joint density of all elements in one column j is equal to the product of the density of individual element:

$$p(\tilde{\gamma}_{jj}, ..., \gamma_{nj} \,|\, y^{\star}, s, \tilde{h}) = \prod_{i=j}^{n} p(\tilde{\gamma}_{ij}, \,|\, y^{\star}, s, \tilde{h}) \tag{3.27}$$

Since all densities on the right hand side are univariate Gaussian densities, these can easily be evaluated. The density $p(\tilde{\gamma}_{jj}, ..., \gamma_{nj} \,|\, y)$ that is marginal of model elements is then obtained as the Monte Carlo average over $p(\tilde{\gamma}_{jj}, ..., \gamma_{nj} \,|\, y^{\star}, s, h)$ using the posterior draws for s, $\tilde{h}$ and all other model parameters in case of more complex models.

3.C   Additional Figures

Figure 3.C.1: Posterior densities of the scales and normalized loading in the CFSV and ISV models: Simple DGP
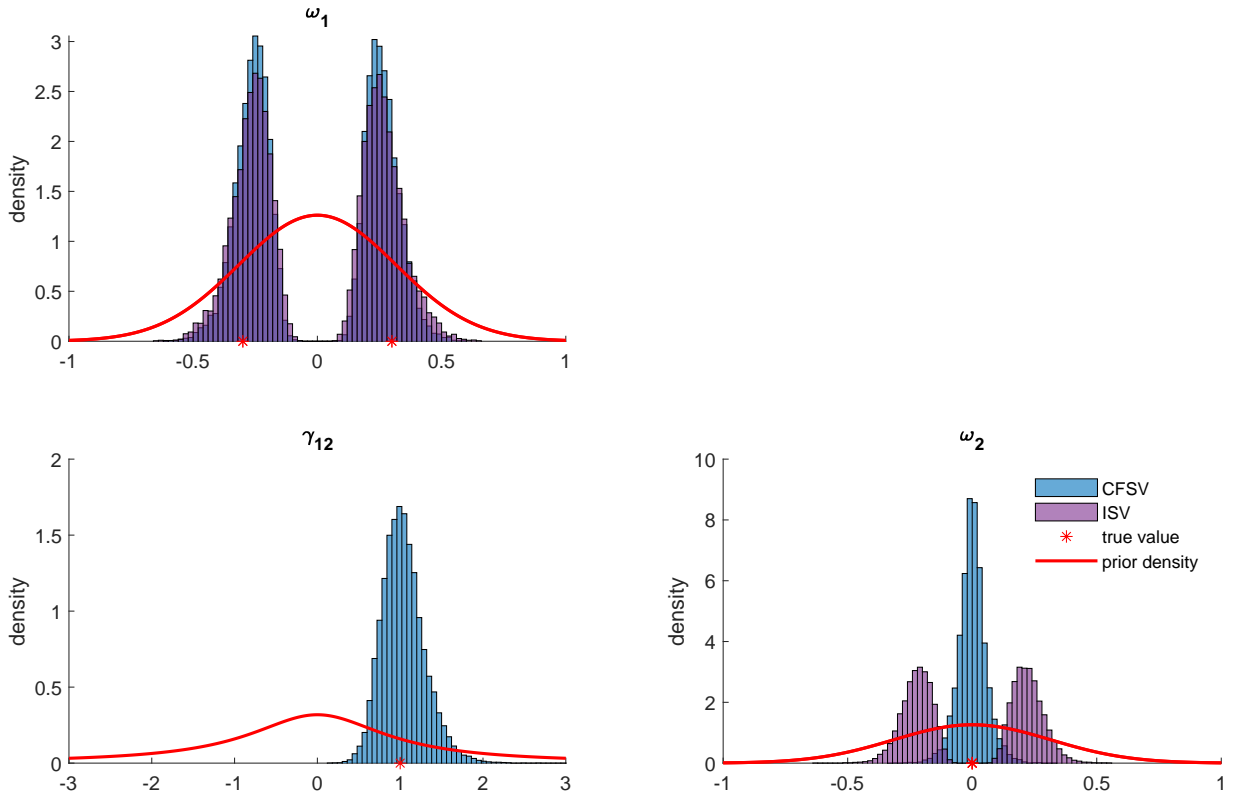
Notes: The normalized loading γ_{12} is obtained as $\frac{\tilde{\gamma}_{12}}{\tilde{\gamma}_{11}}$ and the scales are defined as $\omega_1 = \tilde{\gamma}_{11}$ and $\omega_2 = \tilde{\gamma}_{22}$. The priors are Gaussian densities for ω_1 and ω_2 and a t_1 density for γ_{12}.

References

Andrade, P., J. Beckenfelder, F. De Fiore, and P. Karadi (2016). The ECB's Asset Purchase Programme: An Early Assessment. Working Paper 1956, Eurpean central Bank.

Antolín-Díaz, J., I. Petrella, and J. F. Rubio-Ramírez (2021). Structural Scenario Analysis with SVARs. *Journal of Monetary Economics 117*, 798–815.

Antolín-Díaz, J. and J. F. Rubio-Ramírez (2018). Narrative Sign Restrictions for SVARs. *American Economic Review 108*(10), 2802–2829.

Arias, J. E., D. Caldara, and J. F. Rubio-Ramírez (2019). The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure. *Journal of Monetary Economics 101*, 1–13.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018). Inference Based on Structural Vector Autoregressions Identified With Sign and Zero Restrictions: Theory and Applications. *Econometrica 86*(2), 685–720.

Aruoba, S. B. and F. Schorfheide (2011). Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-Offs. *American Economic Journal: Macroeconomics 3*(1), 60–90.

Bachmann, R. and E. R. Sims (2012). Confidence and the Transmission of Government Spending Shocks. *Journal of Monetary Economics 59*(3), 235–249.

Bańbura, M. and A. van Vlodrop (2018). Forecasting with Bayesian Vector Autoregressions with Time Variation in the Mean. Tinbergen Institute Discussion Papers 18-025/IV, Tinbergen Institute.

Baumeister, C. and J. D. Hamilton (2015). Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *Econometrica 83*(5), 1963–1999.

Belongia, M. T. and P. N. Ireland (2016). Money and Output: Friedman and Schwartz Revisited. *Journal of Money, Credit and Banking 48*(6), 1223–1266.

Bernanke, B. S. (2007). Inflation Expectations and Inflation Forecasting. Speech, Monetary Economics Workshop of the National Bureau of Economic Research.

Bertsche, D. and R. Braun (2022). Identification of Structural Vector Autoregressions by

References

Stochastic Volatility. *Journal of Business & Economic Statistics 40*(1), 328–341.

Bobeica, E., M. Ciccarelli, and I. Vansteenkiste (2019). The Link between Labor Cost and Price Inflation in the Euro Area. Working Paper 2235, European Central Bank.

Brunnermeier, M., D. Palia, K. A. Sastry, and C. A. Sims (2021). Feedbacks: Financial Markets and Economic Activity. *American Economic Review 111*(6), 1845–79.

Caldara, D. and E. Herbst (2019). Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. *American Economic Journal: Macroeconomics 11*(1), 157–192.

Caldara, D. and C. Kamps (2017). Analytics of SVARs: A Unified Framework to Measure Fiscal Multiplier. *Review of Economic Studies 84*, 1015–1040.

Canova, F. and L. Gambetti (2010). Do Expectations Matter? The Great Moderation Revisited. *American Economic Journal: Macroeconomics 2*(3), 183–205.

Carriero, A., T. E. Clark, and M. Marcellino (2016). Common Drifting Volatility in Large Bayesian VARs. *Journal of Business & Economic Statistics 34*(3), 375–390.

Carriero, A., T. E. Clark, and M. Marcellino (2018). Measuring Uncertainty and its Impact on the Economy. *Review of Economics and Statistics 100*(5), 799–815.

Carriero, A., T. E. Clark, and M. Marcellino (2019). Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors. *Journal of Econometrics 212*(1), 137–154.

Carriero, A., F. Corsello, and M. Marcellino (2022). The Global Component of Inflation Volatility. *Journal of Applied Econometrics 37*(4), 700–721.

Carvalho, C., S. Eusepi, E. Mönch, and B. Preston (2022). Anchored Inflation Expectations. *American Economic Journal: Macroeceonomics* (forthcoming).

Carvalho, V. M. and A. C. Harvey (2005). Convergence in the Trends and Cycles of Euro-Zone Income. *Journal of Applied Econometrics 20*(2), 275–289.

Castelnuovo, E., L. Greco, and D. Raggi (2014). Policy Rules, Regime Switches, and Trend Inflation: An Empirical Investigation for the United States. *Macroeconomic Dynamics 18*(4), 920–942.

Chan, J. C. (2018). Specification Tests for Time-Varying Parameter Models with Stochastic Volatility. *Econometric Reviews 37*(8), 807–823.

Chan, J. C. (2022). Comparing Stochastic Volatility Specifications for Large Bayesian VARs. Preprint arxiv:2208.13255.

Chan, J. C. and E. Eisenstat (2015). Marginal Likelihood Estimation with the Cross-Entropy Method. *Econometric Reviews 34*(3), 256–285.

References

Chan, J. C. and E. Eisenstat (2018). Bayesian Model Comparison for Time-Varying Parameter VARs with Stochastic Volatility. *Journal of Applied Econometrics 33*(4), 509–532.

Chan, J. C. and A. L. Grant (2015). Pitfalls of Estimating the Marginal Likelihood using the Modified Harmonic Mean. *Economics Letters 131*, 29–33.

Chan, J. C. and A. L. Grant (2016). Fast Computation of the Deviance Information Criterion for Latent Variable models. *Computational Statistics and Data Analysis 100*, 847–859.

Chan, J. C. and I. Jeliazkov (2009). Efficient Simulation and Integrated Likelihood Estimation in State Space Models. *International Journal of Mathematical Modelling and Numerical Optimisation 1*(1-2), 101–120.

Chan, J. C., G. Koop, and X. Yu (2022). Large Order-Invariant Bayesian Vars with Stochastic Volatility.

Chan, J. C. C., T. E. Clark, and G. Koop (2018). A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations. *Journal of Money, Credit and Banking 50*(1), 5–53.

Ciccarelli, M., J. A. Garcia, and C. Montes-Galdon (2017). Unconventional Monetary Policy and the Anchoring of Inflation Expectations. ECB Working Paper 1995.

Clark, T. E. and T. Davig (2011). Decomposing the Declining Volatility of Long-Term Inflation Expectations. *Journal of Economic Dynamics and Control 35*(7), 981–999.

Clark, T. E. and F. Ravazzolo (2015). Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility. *Journal of Applied Econometrics 30*(4), 551–575.

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010). Inflation-Gap Persistence in the US. *American Economic Journal: Macroeconomics 2*(1), 43–69.

Coibion, O. and Y. Gorodnichenko (2011). Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation. *American Economic Review 101*(1), 341–70.

Coibion, O., Y. Gorodnichenko, and R. Kamdar (2018). The Formation of Expectations, Inflation, and the Phillips Curve. *Journal of Economic Literature 56*(4), 1447–91.

Coibion, O., Y. Gorodnichenko, S. Kumar, and M. Pedemonte (2020). Inflation Expectations as a Policy Tool? *Journal of International Economics 124*, 103297.

Crump, R. K., S. Eusepi, and E. Moench (2018). The Term Structure of Expectations and Bond Yields. Technical Report 775, Federal Reserve Bank of New York.

Cukierman, A. and A. H. Meltzer (1986). A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information. *Econometrica 54*(5), 1099–1128.

References

D'Amico, S. and T. B. King (2017). What Does Anticipated Monetary Policy Do? Working Paper 2015-10, Federal Reserve Bank of Chicago.

Debortoli, D., J. Galí, and L. Gambetti (2020). On the Empirical (Ir)Relevance of the Zero Lower Bound Constraint. *NBER Macroeconomics Annual 34*(1), 141–170.

Del Negro, M. and S. Eusepi (2011). Fitting Observed Inflation Expectations. *Journal of Economic Dynamics and Control 35*(12), 2105–2131.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017). Safety, Liquidity, and the Natural Rate of Interest. *Brookings Papers on Economic Activity 48*(1 (Spring)), 235–316.

Diegel, M. and D. Nautz (2021). Long-Term Inflation Expectations and the Transmission of Monetary Policy Shocks: Evidence from a SVAR Analysis. *Journal of Economic Dynamics and Control 130*, 104192.

Doh, T. and A. Oksol (2018). Has the Anchoring of Inflation Expectations Changed in the United States during the Past Decade? *Economic Review* (QI), 31–58.

ECB (2021). ECB's Governing Council Approves its New Monetary Policy Strategy. Press Release, European Central Bank.

Erceg, C. J. and A. T. Levin (2003). Imperfect Credibility and Inflation Persistence. *Journal of Monetary Economics 50*(4), 915–944.

Frühwirth-Schnatter, S. and H. Wagner (2010). Stochastic Model Specification Search for Gaussian and Partial Non-Gaussian State Space Models. *Journal of Econometrics 154*(1), 85–100.

Fry, R. and A. Pagan (2011). Sign Restrictions in Structural Vector Autoregressions: A Critical Review. *Journal of Economic Literature 49*(4), 938–960.

Fuhrer, J. (2012). The Role of Expectations inInflation Dynamics. *International Journal of Central Banking 8*(1).

Gambetti, L. and A. Musso (2020). The Macroeconomic Impact of the ECB's Expanded Asset Purchase Programme. *European Economic Review 130*.

Gáti, L. (2022). Monetary policy and Anchored Expectations: An Endogenous Gain Learning Model. Working Paper 2685, European Central Bank.

Geiger, M. and J. Scharler (2021). How Do People Interpret Macroeconomic Shocks? Evidence from US Survey Data. *Journal of Money, Credit and Banking 53*(4), 813–843.

Gertler, M. and P. Karadi (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Giacomini, R. and T. Kitagawa (2021). Robust Bayesian Inference for Set-Identified Models. *Econometrica 89*(4), 1519–1556.

References

Goodfriend, M. (2004). Inflation Targeting in the United States? In *The inflation-targeting debate*, pp. 311–352. University of Chicago Press.

Goodfriend, M. (2007). How the World Achieved Consensus on Monetary Policy. *Journal of Economic Perspectives 21*(4), 47–68.

Grant, A. L. and J. C. Chan (2017). A Bayesian Model Comparison for Trend-Cycle Decompositions of Output. *Journal of Money, Credit and Banking 49*(2-3), 525–552.

Granziera, E., H. R. Moon, and F. Schorfheide (2018). Inference for VARs Identified with Sign Restrictions. *Quantitative Economics 9*(3), 1087–1121.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005). The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models. *American Economic Review 95*(1), 425–436.

Hachula, M. and D. Nautz (2018). The Dynamic Impact of Macroeconomic News on Long-Term Inflation Expectations. *Economics Letters 165*, 39–43.

Inoue, A. and L. Kilian (2020a). Joint Bayesian Inference about Impulse Responses in VAR Models. Working Papers 2022, Federal Reserve Bank of Dallas.

Inoue, A. and L. Kilian (2020b). The Role of the Prior in Estimating VAR models with Sign Restrictions. CEPR Discussion Paper No. DP15545.

Ireland, P. N. (2007). Changes in the Federal Reserve's Inflation Target: Causes and Consequences. *Journal of Money, Credit and Banking 39*(8), 1851–1882.

Jarociński, M. and P. Karadi (2020). Deconstructing Monetary Policy Surprises - The Role of Information Shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Jorgensen, P. L. and K. J. Lansing (2022). Anchored Inflation Expectations and the Slope of the Phillips Curve. Working Paper 2019-27, Federal Reserve Bank of San Francisco.

Kass, R. E. and A. E. Raftery (1995). Bayes Factors. *Journal of the american statistical association 90*(430), 773–795.

Kastner, G. (2019). Sparse Bayesian Time-Varying Covariance Estimation in Many Dimensions. *Journal of Econometrics 210*(1), 98–115.

Kastner, G., S. Frühwirth-Schnatter, and H. F. Lopes (2017). Efficient Bayesian Inference for Multivariate Factor Stochastic Volatility Models. *Journal of Computational and Graphical Statistics 26*(4), 905–917.

Kilian, L. and L. T. Lewis (2011). Does the Fed Respond to Oil Price Shocks? *The Economic Journal 121*(555), 1047–1072.

Kilian, L. and H. Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press.

References

Kim, C.-J. and J. Kim (2022). Trend-Cycle Decompositions of Real GDP Revisited: Classical and Bayesian Perspectives on an Unsolved Puzzle. *Macroeconomic Dynamics 26*(2), 394–418.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The review of economic studies 65*(3), 361–393.

Kozicki, S. and P. A. Tinsley (2005). Permanent and Transitory Policy Shocks in an Empirical Macro Model with Asymmetric Information. *Journal of Economic Dynamics and Control 29*(11), 1985–2015.

Krippner, L. (2013). Measuring the Stance of Monetary Policy in Zero Lower Bound Environments. *Economics Letters 118*(1), 135–138.

Kroese, D. P., J. C. Chan, et al. (2014). *Statistical Modeling and Computation.* Springer.

Lansing, K. J. (2009). Time-Varying US Inflation Dynamics and the New Keynesian Phillips Curve. *Review of Economic Dynamics 12*(2), 304–326.

Leduc, S., K. Sill, and T. Stark (2007). Self-fulfilling Expectations and the Inflation of the 1970s: Evidence from the Livingston Survey. *Journal of Monetary Economics 54*(2), 433–459.

Lindsey, D. E., A. Orphanides, R. H. Rasche, et al. (2013). The Reform of October 1979: How it happened and Why. *Federal Reserve Bank of St. Louis Review 95*(6), 487–542.

Lombardi, M. and F. Zhu (2019). A Shadow Policy Rate to Calibrate U.S. Monetary Policy at the Zero Lower Bound. *International Journal of Central Banking*, 305–346.

Mertens, E. (2016). Measuring the Level and Uncertainty of Trend Inflation. *The Review of Economics and Statistics 98*(5), 950–967.

Mertens, E. and J. M. Nason (2020). Inflation and Professional Forecast Dynamics: An Evaluation of Stickiness, Persistence, and Volatility. *Quantitative Economics 11*(4), 1485–1520.

Milani, F. (2020). Learning and the Evolution of the Fed's Inflation Target. *Macroeconomic Dynamics 24*(8), 1904–1923.

Nautz, D., L. Pagenhardt, and T. Strohsal (2017). The (De-)Anchoring of Inflation Expectations: New Evidence from the Euro Area. *The North American Journal of Economics and Finance 40*, 103–115.

Nautz, D., T. Strohsal, and A. Netšunajev (2019). The Anchoring of Inflation Expectations in the Long and in the Short Run. *Macroeconomic Dynamics 23*(5), 1959–1977.

Powell, J. H. (2020). New Economic Challenges and the Fed's Monetary Policy Review. Speech, Federal Reserve Bank of Kansas City, Jackson Hole.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary

Policy. *The Review of Economic Studies 72*(3), 821–852.

Schnabel, I. (2022). Monetary Policy and the Great Volatility. Speech, Federal Reserve Bank of Kansas City, Jackson Hole.

Shapiro, A. H. and D. Wilson (2019, June). Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analytics. Technical Report 2019-02, Federal Reserve Bank of San Francisco.

Sims, C. A. and T. Zha (2006a). Does Monetary Policy Generate Recessions? *Macroeconomic Dynamics 10*(2), 231–272.

Sims, C. A. and T. Zha (2006b). Were There Regime Switches in US Monetary Policy? *American Economic Review 96*(1), 54–81.

Smets, F. and R. Wouters (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American economic review 97*(3), 586–606.

Uhlig, H. (2005). What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure. *Journal of Monetary Economics 52*(2), 381–419.

Verdinelli, I. and L. Wasserman (1995). Computing Bayes factors using a Generalization of the Savage-Dickey Density Ratio. *Journal of the American Statistical Association 90*(430), 614–618.

Watson, M. (2019). Comment on 'On the Empirical (Ir)Relevance of the Zero Lower Bound' by D. Debortoli, J. Gali, and L. Gambetti . *NBER Macroeconomics Annual 34*, 182–193.

Wolf, C. K. (2020). SVAR (Mis-)Identification and the Real Effects of Monetary Policy Shocks. *American Economic Journal: Macroeconomics 12*(4), 1–32.

Wong, B. (2015). Do Inflation Expectations Propagate the Inflationary Impact of Real Oil Price Shocks?: Evidence from the Michigan Survey. *Journal of Money, Credit and Banking 47*(8), 1673–1689.

Wu, J. C. and F. D. Xia (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking 48*(2-3), 253–291.

List of Publications

Publications in peer-reviewed Journals

Diegel, M., and Nautz, D. (2021). Long-term Inflation Expectations and the Transmission of Monetary Policy Shocks: Evidence from a SVAR analysis. *Journal of Economic Dynamics and Control (130)*:104192.
Publication of Chapter 1.

Publications as Working Papers

Diegel, M., and Nautz, D. (2020). The Role of Long-Term Inflation Expectations for the Transmission of Monetary Policy Shocks. Free University Berlin, School of Business & Economics. Diskussionsbeiträge No. 2020/19.
Preliminary version of Chapter 1.

Zusammenfassung

In modernen makroökonomischen Modellen sind die Inflationserwartungen eine wichtige Determinante der tatsächlichen Inflationsrate. Daher sind Zentralbanken ihr Mandadt der Preisstabilität im Sinne eines konkreten Inflationszielen verstehen um eine feste Verankerung der längerfristigen Inflationserwartungen an ihrem Ziel bemüht. In dieser Arbeit wird die Verankerung der langfristigen Inflationserwartungen in den USA aus drei sich ergänzenden Perspektiven untersucht.

Im ersten Kapitel, das in Zusammenarbeit mit Dieter Nautz entstanden ist, wird die Rolle der langfristigen Inflationserwartungen für die geldpolitische Transmission in den USA in einer strukturellen Vektorautoregression (SVAR) untersucht. Im Gegensatz zu früheren Studien stellen wir fest, dass die langfristigen Inflationserwartungen in den USA unmittelbar auf einen geldpolitischen Schock reagieren. Die Analyse struktureller Szenarien zeigt erstens, dass die langfristigen Inflationserwartungen, im Einklang mit einem Wiederverankerungskanal der Geldpolitik, eine wichtige Rolle bei der Übertragung von geldpolitischen Schocks auf die tatsächliche Inflationsrate spielen. Ein zweites Szenario legt nahe, dass die Reaktion der Geldpolitik auf Erwartungsschocks zur Stabilisierung von Inflation und Arbeitslosigkeit während der ersten Nullzinsphase nach der Finanzkrise von 2008 beigetragen hat.

Das zweite Kapitel analysiert die zeitlich variierende Glaubwürdigkeit des Inflationsziels der US-amerikanischen Zentralbank, der Federal Reserve Bank (Fed), in einem empirischen Makromodell mit asymmetrischer Information, bei dem die Öffentlichkeit über das tatsächliche Inflationsziel aus der Zinspolitik der Zentralbank lernt. Um die sich weiterentwickelnde Kommunikationsstrategie der Fed durch die verschiedenen geldpolitischen Regime zu erfassen, lässt die Schätzung Brüche in der Lernregel und den Varianzen der strukturellen Schocks zu. Niedrige Glaubwürdigkeit tritt während der Volcker Disinflation und in geringerem Maße nach der Finanzkrise 2008 auf. Die Ankündigung des 2-Prozent Inflationsziels im Jahr 2012 hatte keinen großen Einfluss auf die Lernregel, verringerte aber die Varianz der transitorischen geldpolitischen Schocks und trägt so zu einer besseren Verankerung des

wahrgenommenen Inflationsziels der Öffentlichkeit bei. Die Ergebnisse legen außerdem nahe, dass die langfristigen Inflationserwartungen der Fachleute nicht mit dem wahrgenommenen Inflationsziel der Öffentlichkeit gleichgesetzt werden sollten.

Im dritten Kapitel werden die Auswirkungen der Volatilität auf die Inflationserwartungen einer SVAR mit stochastischer Volatilität (SV) untersucht. Die überwiegende Mehrheit der SVAR-SV Literatur macht jedoch die Annahme, dass die Volatilitäten aller Schocks sich *unabhängig* voneinander entwickeln und übersieht so mögliche gemeinsame Quellen der Zeitvariation in den Volatilitäten. Daher wird im dritten Kapitel vorgeschlagen, gemeinsame Faktoren in der stochastischen Volatilität (CFSV) zuzulassen. Unter Verwendung dieses neuen Volatilitätsmodells wird die Volatilität von kurz- und langfristigen Inflationserwartungen im Modell von Clark und Davig (2011) erneut untersucht. Die Ergebnisse zeigen, dass, um die zeitliche Variation der Volatilitäten des fünfdimensionalen SVAR zu erfassen, nur zwei Faktoren erforderlich sind. Während der erste Faktor ausschließlich die Volatilität des eigenen Schocks der langfristigen Erwartungen bestimmt, beeinflusst der zweite Faktor sowohl die Volatilität der Schocks der tatsächlichen Inflationsrate als auch die der kurzfristigen Inflationserwartungen. Dies hat möglicherweise Auswirkungen auf die Verankerung der Inflationserwartungen, insbesondere in Zeiten erhöhter Inflationsvolatilität.

Summary

Inflation expectations are a key determinant of actual inflation in modern macroeconomic models. Therefore, inflation targeting central banks are concerned with a firm anchoring of inflation expectations, especially longer-term expectations. This dissertation investigates the anchoring of US long-term inflation expectations from three complementary perspectives.

The first chapter, which is joint work with Dieter Nautz, investigates the role of long-term inflation expectations for the monetary transmission mechanism and the conduct of monetary policy in a structural VAR framework. In contrast to earlier studies, we find that US long-term inflation expectations respond significantly to a monetary policy shock. A first structural scenario analysis suggests that, in line with a re-anchoring channel of monetary policy, long-term inflation expectations play an important role for the transmission of monetary policy shocks to the rate of inflation. A second scenario shows that the response of monetary policy to expectations shocks has contributed to the stabilization of inflation and unemployment when the zero lower bound was binding after the 2008 Financial Crisis.

The second chapter analyzes the time-varying credibility of the Fed's inflation target in an empirical macro model with asymmetric information. Due to asymmetric information, the public has to learn about the actual inflation target from the Fed's interest rate policy. To capture the evolving communication strategy of the Fed, the learning rule and the structural shock variances are allowed to change across monetary policy regimes. I find that imperfect credibility is pronounced during the Volcker Disinflation and to a lesser extent in the aftermath of the 2008 Financial Crisis. The announcement of the 2% inflation target in 2012 did not affect the learning rule strongly but reduced the variance of transitory monetary policy shocks, which improves the degree of anchoring of the perceived target to the actual inflation target. Furthermore, the results caution against equating long-term inflation expectations of professionals with the perceived inflation target.

The third chapter examines volatility spill-overs to inflation expectation in a structural vector autoregressive (SVAR) model with stochastic volatility (SV). However, the vast majority of the SVAR-SV literature assumes that the volatilities evolve *independently* and,

thus, ignores possible common sources of the time-variation in volatilities. To overcome this problem, the third chapter proposes a new model that allows for common factors in the stochastic volatility (CFSV). With this new model I revisit the application of Clark and Davig (2011), who decompose the decline in the volatility of short- and long-term inflation expectations. Only two factors are required to capture the time variation of the five-dimensional SVAR. While the first factor is exclusive to the volatility of the own shock to long-term expectations, the second factor loads onto both the volatility of shocks to actual inflation and short-term inflation expectations. This has implications for the anchoring of inflation expectations, especially in times of increased inflation volatility .

Erklärungen

Erklärung gem. §4 PO

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, den 14. Oktober 2022

....

Max Diegel

Erklärung gem. § 10 Abs. 3

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe:

- MATLAB

- Python

- Excel

- Latex

- Siehe References, für die verwendete Literatur

Auf dieser Grundlage habe ich die Arbeit selbstständig verfasst.

Berlin, den 14. Oktober 2022

....
Max Diegel