Robust Real Rate Rules

Tom D. Holden, Deutsche Bundesbank[1] 28/02/2022

Abstract: Central banks would like to ensure determinate inflation, to rule out self-fulfilling fluctuations. Traditional monetary rules can fail to produce determinacy under many conditions. This paper proposes a family of monetary rules that ensure determinate inflation under the weakest possible assumptions about the behaviour of households and firms. Despite this, the family of rules is general enough to allow the determinate implementation of arbitrary inflation dynamics. The rules are easy to implement in practice, and even simple rules in our class attain high welfare. The success of these rules provides new insight into the core monetary transmission mechanism. The Fisher equation is key, not the Euler equation or the Phillips curve.

TODO Mention other questions: Monetary policy with long bonds, irregular meetings, Implementing optimal policy, Why Phillips not useful for forecasting, Redistribution channel of monetary policy

Keywords: robust monetary rules, determinacy, Taylor principle, inflation dynamics, monetary transmission mechanism

JEL codes: E52, E43, E31

---

[1] Wilhelm-Epstein-Straße 14, 60431, Frankfurt am Main, Germany. E-mail: thomas.holden@gmail.com.

The views expressed in this paper are those of the author and do not represent the views of the Deutsche Bundesbank, the Eurosystem or its staff.

The author would like to thank Nicola Borri, Xiaoshan Chen, Felix Geiger, Refet Gürkaynak, Peter Ireland, Malte Knueppel, Philipp Lieberknecht, Elmar Mertens, Stéphane Moyen, Lars Svensson and Mu-Chun (Tomi) Wang for helpful discussions.

This paper incorporates results from an earlier note, circulated under the title "The One Equation New Keynesian Model". An early draft of this paper was previously circulated under the title "A Robust Monetary Rule".

Today you start work as president of the Fictian Central Bank (FCB). As FCB president, you have a clear mandate to stabilize inflation, even if that results in unemployment or output losses. How should you act? You have studied New Keynesian macro, so you are inclined to follow some variant of the Taylor rule. You recall the prescription of the Taylor principle: the response of nominal rates to inflation should be greater than one to ensure determinacy and rule out self-fulfilling fluctuations in inflation. But you also remember reading other papers which talked of the Taylor principle being insufficient if there are hand-to-mouth households (Gali, Lopez-Salido & Valles 2004), firm-specific capital (Sveen & Weinke 2005), high government spending (Natvik 2009), or if the inflation target is positive (Ascari & Ropele 2009), particularly in the presence of trend growth and sticky wages (Khan, Phaneuf & Victor 2019). Indeed, you recollect that the Taylor principle inverts if there are sufficiently many hand-to-mouth households (Bilbiie 2008), certain financial frictions (Manea 2019), or non-rational expectations (Branch & McGough 2010; 2018). Is there a way you could act to ensure determinacy even if one or more of these circumstances is true? This paper provides a family of "robust real rate rules" that manage to do this.

To illustrate the idea behind these rules, suppose that both nominal and real bonds are traded in an economy. If a unit of the former is purchased at t, it returns the principal plus a nominal yield of i_t in period $t+1$. If a unit of the latter is purchased at t, it returns the principal plus a nominal yield of $r_t + \pi_{t+1}$ in period $t+1$, where π_{t+1} is realized inflation between t and $t+1$. Abstracting for the moment from inflation risk premia, term premia and liquidity premia, arbitrage between these two markets implies that the Fisher equation must hold, i.e.:

$$i_t = r_t + \mathbb{E}_t \pi_{t+1}, \tag{1}$$

where $\mathbb{E}_t \pi_{t+1}$ is the full information rational expectation of period $t+1$'s inflation rate, given period t's information. Suppose further than the central bank observes both the nominal and real bond markets, and that it can intervene in the former. Then the central bank can choose to set nominal interest rates according to the simple rule:[2]

$$i_t = r_t + \phi\pi_t, \tag{2}$$

where $\phi > 1$. Combining these two equations gives that:

$$\mathbb{E}_t\pi_{t+1} = \phi\pi_t,$$

which has a unique non-explosive solution of $\pi_t = 0$.[3] Determinate inflation!

Why is this robust? Firstly, the rule does not require the aggregate Euler equation to hold, even approximately. For the Fisher equation (1) to hold (still ignoring risk/term/liquidity premia for now), there only need to be two deep pocketed, fully informed, rational agents. Arbitrage takes care of the rest. Even full information is not necessary. Since large markets aggregate information (Hellwig 1980; Lou et al. 2019), the Fisher equation can come to hold even when information about future inflation is dispersed amongst market participants.

Given that the rule does not require the aggregate Euler equation to hold, it is automatically robust to heterogeneity, hand-to-mouth agents and non-rational consumer expectations. The only expectations that matter are the expectations of participants in the markets for nominal and real bonds. It is surely much more

---

[2] Such rules have appeared in Adão, Correia & Teles (2011), Holden TODO UPDATE (2019) and Lubik, Matthes & Mertens (2019) amongst other places. However, in the prior literature they have chiefly been introduced for analytic convenience, rather than as serious proposals. The one exception is Cochrane (2017) who briefly discusses a rule of this form before moving on to discuss rules which hold $i_t - r_t$ constant (i.e. rules with $\phi = 0$). Cochrane (2018) further explores rules holding $i_t - r_t$ constant. The "indexed payment on reserve" rules of Hall & Reis (2016) also rely on observable real rates, but use a different mechanism to achieve determinacy. They propose that the CB issues an asset ("reserves") with nominal return from \$1 of $\$(1 + r_t)\frac{p_{t+1}}{p_t^*}$ or $\$(1 + i_t)\frac{p_t}{p_t^*}$.

Additionally, in older work, Hetzel (1990) proposes using the spread between nominal and real bonds to guide monetary policy, and Dowd (1994) proposes targeting the price of futures contracts on the price level, which has a similar flavour to our rules, since our rules effectively use expected inflation as the instrument of monetary policy.

There is also an established literature looking at rules tracking the efficient real interest rate, see e.g. Cúrdia et al. (2015), which is a very different idea.

[3] Here we sidestep the issues raised by Cochrane (2011) and follow the standard New Keynesian literature in assuming agents will always select non-explosive paths for inflation. The escape clause rules of Christiano & Takahashi (2018) are one way by which central banks could ensure coordination on the expectations consistent with non-explosive inflation.

reasonable to assume that financial market outcomes lead to rational expectations than to assume rationality of households more generally.

Secondly, the rule does not require the aggregate Phillips equation to hold. The slope of the Phillips curve will have no impact on the dynamics of inflation. If the FCB president is unconcerned with output, they do not need to know if the Phillips curve holds, let alone its slope. Nor does it matter how firms form inflation expectations. Inflation is pinned down by the Fisher and monetary rules, so while non-rational firm expectations could affect output fluctuations, they will not alter the dynamics of inflation.

This may be surprising. How could price setters fail to determine inflation? The short answer is "Walras's law". To see how this plays out, suppose that today all firms decide to double their price. Financial market participants still expect zero inflation next period, because that is the only outcome consistent with non-explosive inflation in future. Thus, financial market participants always value nominal bonds the same as real bonds. But the central bank's monetary rule instructs it to attempt to produce nominal rates which are much higher than real rates, as today's inflation is high. So, the central bank wants to sell nominal bonds, i.e., to borrow money from financial market participants.

However, no amount of nominal bond selling will induce market participants to lower their valuation of nominal bonds below that of real bonds, though both valuations may fall together (i.e., both nominal and real rates rise). Thus, the central bank will end up reducing the money supply to zero. With households having zero cash, not all final goods will be sold.[4] Thus, the final goods market will not clear. To obtain market clearing in final goods, at least some price setters must reduce their price until inflation is zero, so ensuring that the central bank sets nominal rates equal to real rates.

[4] For example, if there are cash goods and credit goods, only credit goods will be sold.

The rest of this paper further examines such "real rate rules". The next section generalizes the simple rule of equation (2) along various dimensions, including examining rules that respond to other endogenous variables. Section 1 goes on to show that there are similar rules that determinately implement an arbitrary path for inflation, robustly across models. Section 2 discusses how a real rate rule could be implemented in practice. We show that it is easy to adapt real rate rules to work with longer bonds.

In the following section (3), we examine the performance of these rules in a tractable model of the redistribution channel of monetary policy. This serves two main purposes. Firstly, it shows robustness to a monetary transmission mechanism that exists even under flexible prices, and which is present in many heterogeneous agent models. Secondly, the model is so simple that it has an exact analytic solution even in the presence of stochastic volatility. We show that time varying risk premia are not a substantial obstacle to the use of these real rate rules. Additionally, we are able to use this simple analytic framework to reassess the costs of indeterminacy. Unlike in linearized models, indeterminacy leads to increases in the mean of inflation, not just its variance.

Next, Section 4 looks at the welfare costs of various real rate rule specifications in New Keynesian models. Section 5 looks at the data to see how close existing central bank behaviour is to following a rule of our proposed form. Finally, Section 6 looks at the consequences of the zero lower bound for the performance of these rules.

1 Generalizations and generality

While the simple rule (2) always produces zero inflation, slight extensions of the rule allow inflation to move. For example, we may add a monetary policy shock, ζ_t to the rule, giving:

$$i_t = r_t + \phi\pi_t + \zeta_t. \tag{3}$$

Monetary policy shocks may perhaps reflect the central bank's limited information. If the central bank does not perfectly observe current inflation, and sets interest rates to $i_t = r_t + \phi\tilde{\pi}_t$, where $\tilde{\pi}_t$ is its signal about inflation, then it will end up setting a slightly different level for nominal rates than that dictated by the rule $i_t = r_t + \phi\pi_t$, effectively generating monetary policy shocks.[5]

The central bank might also deliberately decide to introduce monetary policy shocks correlated with the economy's structural shocks. For example, by lowering $i_t - r_t$ following a positive mark-up or cost-push shock, the central bank can lessen the movement in the output gap.[6] This has no effect on the determinacy region as structural shocks are exogenous. For now though, we assume that ζ_t is independent of other structural shocks.

From combining (3) with the Fisher equation (1) we have:

$$\mathbb{E}_t\pi_{t+1} = \phi\pi_t + \zeta_t,$$

which (with $\phi > 1$) has the unique solution $\pi_t = -\frac{1}{\phi-\rho_\zeta}\zeta_t$, if ζ_t follows an AR(1) process with persistence ρ_ζ.

A contractionary (positive) monetary policy shock results in a fall in inflation, as expected. If the central bank is more aggressive, so ϕ is larger, then inflation is less volatile. Only monetary policy shocks affect inflation. Of course, if there is a nominal rigidity in the model, such as sticky prices or wages, monetary shocks may have an impact on real variables. But as long as the central bank follows rules like this, these real disruptions have no feedback to inflation. We can understand inflation without worrying about the rest of the economy.

---

[5] Lubik, Matthes & Mertens (2019) look at the determinacy consequences of a central bank that filters inflation signals in order to retrieve the optimal estimate. The determinacy problems they highlight all disappear if the central bank directly responds to its signal.

[6] Ireland (2007) presents evidence that the US Federal Reserve has reacted to mark-up shocks.

In line with this, an extensive body of empirical evidence finds no role for the Phillips curve in forecasting inflation (see e.g. Atkeson & Ohanian 2001; Ang, Bekaert & Wei 2007; Stock & Watson 2009; Dotsey, Fujita & Stark 2018). In a recent contribution, Dotsey, Fujita & Stark (2018) find that in the post-1984 period, Phillips curve based forecasts perform worse than those of a simple IMA(1,1) model, both unconditionally and conditional on various measures of the state of the economy. This provides strong support for models in which the causation in the Phillips curve runs in only one direction: from inflation to the output gap.[7]

1.1 Robust real rate rules in the three equation NK world

To understand how our robust rule in equation (3) can explain causation running from inflation to the output gap in the Phillips curve, suppose the rest of the model comprises the Phillips curve:[8]

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + \kappa \omega_t, \tag{4}$$

and the discounted/compounded Euler equation:

$$x_t = \delta \mathbb{E}_t x_{t+1} - \varsigma (r_t - n_t), \tag{5}$$

where x_t is the output gap, ω_t is a mark-up/cost-push shock, and n_t is the exogenous natural real rate of interest. This form of discounted/compounded Euler equation appears in Bilbiie (2019) and (under discounting) in McKay, Nakamura & Steinsson (2017). The latter paper shows it provides a good approximation to a heterogeneous agent model with incomplete markets. The standard Euler equation is recovered if $\delta = 1$ and ς is the elasticity of intertemporal substitution. This specification also nests the limited asset market participation or "TANK" model of Bilbiie (2008) when $\delta = 1$, but ς is allowed to be negative.

[7] McLeay & Tenreyro (2019) provide an alternative explanation based on the fact that optimal policy prescribes a negative correlation between inflation and output, making difficult empirical identification of the Phillips curve.

[8] Throughout this paper, we multiply the mark-up shock by κ as the ratio of the response to x_t and the response to ω_t is not a function of either the (Calvo) price adjustment probability or the (Rotemberg) price adjustment cost. See Khan (2005) for derivations.

Since $\pi_t = -\frac{1}{\phi-\rho_\zeta}\zeta_t$, and ζ_t is AR(1) with persistence ρ_ζ, the Phillips curve (4) implies that $x_t = -\frac{1}{\kappa}\frac{1-\beta\rho_\zeta}{\phi-\rho_\zeta}\zeta_t - \omega_t$. The Phillips curve is determining the output gap, given the already determined level of inflation. Does x_t help forecast π_t here? Clearly no. $\mathbb{E}_t\pi_{t+1} = -\frac{1}{\phi-\rho_\zeta}\mathbb{E}_t\zeta_{t+1} = -\frac{\rho_\zeta}{\phi-\rho_\zeta}\zeta_t = \rho_\zeta\pi_t$. Once you know π_t, you already have all the information you need to form the optimal forecast of π_{t+1}. The correlation in π_t and x_t provides no extra information.[9]

This model also enables us to show the robustness of our rule's determinacy in practice. Note that with x_t expressed as a linear combination of exogenous variables, there is no need to solve the Euler equation (5) forward, so the degree of discounting (δ) can have no effect on determinacy. Not needing to solve the Euler equation forward also gives robustness to a missing transversality constraint on household assets. For example, if ω_t is independent across time, then the Euler equation implies $r_t = n_t + \frac{1}{\varsigma}\left[\frac{1}{\kappa}\frac{(1-\beta\rho_\zeta)(1-\delta\rho_\zeta)}{\phi-\rho_\zeta}\zeta_t + \omega_t\right]$. This contrasts with the results of Bilbiie (2019) who finds that when $\varsigma > 0$ and $\tilde{\beta} \le 1$, the Taylor principle $(\phi > 1)$ is only sufficient for determinacy in the discounting case $(\delta \le 1)$,[10] and with Bilbiie (2008) who finds that when $\delta = 1$ and $\varsigma < 0$, the Taylor principle $(\phi > 1)$ is neither necessary nor sufficient for determinacy.[11] Under our rule (3), the Taylor principle is necessary and sufficient for determinacy whether there is discounting or compounding, and whether ς is positive or negative (given $\phi \ge 0$).[12]

The rule is also robust to the presence of lags in the Euler or Phillips curve. For example, suppose the Phillips curve and Euler equation are instead given by:

$$\begin{aligned} \pi_t &= \tilde{\beta}(1-\varrho_\pi)\mathbb{E}_t\pi_{t+1} + \tilde{\beta}\varrho_\pi\pi_{t-1} + \kappa x_t + \kappa\omega_t, \\ x_t &= \tilde{\delta}(1-\varrho_x)\mathbb{E}_t x_{t+1} + \tilde{\delta}\varrho_x x_{t-1} - \varsigma(r_t - n_t), \end{aligned} \tag{6}$$

---

[9] This result is robust to generalizing to an ARMA$(1,1)$ process for ζ_t. See Appendix A.1.

[10] See equation (40) of Appendix C.1 of Bilbiie (2019).

[11] See Proposition 7 of Appendix B.1 of Bilbiie (2008).

[12] In Appendix A.2 we prove that this is robust to monetary responses to the real rate which are not exactly equal to 1. This is also a corollary of the more general result prove in Appendix A.4.

where $\tilde{\beta}$ and $\tilde{\delta}$ may not have the same structural interpretation as β and δ (depending on the precise micro-foundation). These equations have no impact on the solution for inflation, which remains $\pi_t = -\frac{1}{\phi-\rho_\zeta}\zeta_t$. Instead, the lag in the Euler equation changes the dynamics of real interest rate, with no impact on inflation or output gaps, while the lag in the Phillips curve affects both output gap and real rate dynamics, with no impact on inflation. For example, if ζ_t's law of motion is given by $\zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_{\zeta,t}$, where $\mathbb{E}_{t-1}\varepsilon_{\zeta,t} = 0$, then:

$$x_t = \frac{1}{\kappa}\frac{1}{\phi - \rho_\zeta}\left[\left(\tilde{\beta}\varrho_\pi - \rho_\zeta\left(1 - \tilde{\beta}(1-\varrho_\pi)\rho_\zeta\right)\right)\zeta_{t-1} - \left(1 - \tilde{\beta}(1-\varrho_\pi)\rho_\zeta\right)\varepsilon_{\zeta,t}\right] - \omega_t.$$

As before, the output gap has a closed form solution in terms of the monetary policy and cost push shocks. Despite appearances, inflation is not a true endogenous state, as it must always equal $-\frac{1}{\phi-\rho_\zeta}\zeta_t$. Monetary policy shocks are still always contractionary, but they only have a transitory impact on the output gap if ϱ_π is around $\frac{\rho_\zeta(1-\beta\rho_\zeta)}{\beta(1-\rho_\zeta^2)}$.

1.2 Responding to other endogenous variables

The original Taylor rule contained a response to output. Even with a unit coefficient on the real interest rate, responding to output will change the determinacy conditions, though it still preserves some robustness. To see this, consider the monetary rule:

$$i_t = r_t + \phi_\pi \pi_t + \phi_x x_t + \zeta_t.$$

Assuming the lag-augmented NK Phillips curve (6) continues to hold, this monetary rule is equivalent to the rule:

$$i_t = r_t + \phi_\pi \pi_t + \kappa^{-1}\phi_x\left[\pi_t - \tilde{\beta}(1-\varrho_\pi)\mathbb{E}_t\pi_{t+1} - \tilde{\beta}\varrho_\pi\pi_{t-1}\right] - \phi_x\omega_t + \zeta_t.$$

(This is produced by using the Phillips curve to substitute out the output gap.) Combined with the Fisher equation, we have that:

$$\mathbb{E}_t\pi_{t+1} = \phi_\pi \pi_t + \kappa^{-1}\phi_x\left[\pi_t - \tilde{\beta}(1-\varrho_\pi)\mathbb{E}_t\pi_{t+1} - \tilde{\beta}\varrho_\pi\pi_{t-1}\right] - \phi_x\omega_t + \zeta_t.$$

This has a determinate solution if the quadratic:

$$\left[1 + \kappa^{-1}\phi_x\tilde{\beta}(1-\varrho_\pi)\right]A^2 - (\phi_\pi + \kappa^{-1}\phi_x)A + \kappa^{-1}\phi_x\tilde{\beta}\varrho_\pi = 0$$

has a unique solution for A inside the unit circle. It is sufficient that the quadratic is positive at $A = -1$ but negative at $A = 1$, which holds if and only if:

$$1 + \kappa^{-1}\phi_x(1 + \tilde{\beta}) + \phi_\pi > 0,$$
$$1 - \kappa^{-1}\phi_x(1 - \tilde{\beta}) - \phi_\pi < 0.$$

So, if $\kappa > 0$, $\phi_x > 0$ and $\tilde{\beta} \in [0,1]$ as expected, then it is sufficient that $\phi_\pi > 1$ as before.[13] This is still considerable robustness. Providing there is something like a Phillips curve linking inflation and the output gap, the standard $\phi_\pi > 1$ condition will be sufficient for determinacy. This would not hold with a more standard monetary rule without a response to real rates: in that case determinacy depends on $\tilde{\delta}$ and ς, as shown by the Bilbiie (2008; 2019) results discussed in the last subsection.

Responding to real rates provides additional robustness even with a response to output as it disconnects the Euler equation from the rest of the model. The only remaining role of the Euler equation is to give a path for real rates, given the already determined paths of output and inflation. The Fisher equation, not the Euler equation is central to monetary policy transmission under real rate rules.

For greater robustness, the central bank can replace the response to the output gap with a response to the cost push shock ω_t. With an appropriate response to ω_t, this is observationally equivalent to responding to the output gap, but ensures determinacy under the standard Taylor principle.

However, it may be hard for the central bank to observe the cost push shock. To get round this, suppose that the central bank knows that a Phillips curve in the form of equation (6) holds. (Our results would generalize to other links between real and nominal variables.) For now, suppose the central bank also knows the coefficients in equation (6). Then the central bank could use a rule of the form:

[13] This is stronger than necessary. The second condition states that $\phi_\pi + \kappa^{-1}\phi_x(1 - \tilde{\beta}) > 1$ so a response to the output gap can substitute for a response to inflation. This condition is identical to that for the standard (purely forward looking) three equation NK model with Taylor type rule found in Woodford (2001).

$$i_t = r_t + \phi_\pi \pi_t + \phi_x \big[x_t - \kappa^{-1} \big[\pi_t - \tilde{\beta}(1 - \varrho_\pi) \mathbb{E}_t \pi_{t+1} - \tilde{\beta} \varrho_\pi \pi_{t-1} \big] \big] + \zeta_t.$$

By equation (6), this implies that:

$$i_t = r_t + \phi_\pi \pi_t - \phi_x \omega_t + \zeta_t,$$

as desired. Of course, the central bank is also unlikely to know the exact coefficients in the Phillips curve. However, we show in Appendix A.3 that the central bank may learn these coefficients in real time, without changing the determinacy conditions, at least under reasonable parameter restrictions.[14]

If the central bank wishes to respond to other endogenous variables, a similar approach should be possible if they are aware of the broad form of the model's structural equations. However, the central bank may legitimately worry about having fundamental misconceptions about how the economy works. They can be reassured though that the Taylor principle will be enough for determinacy if the response to other endogenous variables is small enough, no matter the form of the model's other equations. We prove this in Appendix A.4. This also implies that a precise unit response to real rates is not needed for determinacy. Real rates are just another endogenous variable, so determinacy only requires a response that is sufficiently close to one.

Classic results on determinacy in monetary models can be reinterpreted through this lens. Even if the central bank is not responding to real interest rates, it is still likely to be responding to variables that are highly correlated with them. For example, many models contain an Euler equation of the form:

$$1 = \beta (\exp r_t) \mathbb{E}_t \left(\frac{C_t}{C_{t+1}} \right)^{\frac{1}{\varsigma}},$$

[14] It is sufficient (but not necessary) that $\phi_x \geq 0$, $\phi_\pi \geq 0$, $\kappa \geq 0$, $\tilde{\beta} \in [0,1]$, $\varrho_\pi \in [0,1)$, $\rho_\zeta \in [0,1)$ and $\phi_\pi > \max\left\{ \frac{1}{\tilde{\beta}(1-\varrho_\pi)}, 2(1 - \varrho_\pi), \frac{\phi_x(1+\tilde{\beta})}{\kappa} \right\}$.

where C_t is real consumption per capita and ς is the elasticity of intertemporal substitution. Additionally, in many models, in equilibrium, consumption growth roughly follows an ARMA$(1,1)$ process:

$$g_t := \log\left(\frac{C_t}{C_{t-1}}\right) = (1-\rho_g)g + \rho_g g_{t-1} + \varepsilon_{g,t} + \theta_g \varepsilon_{g,t-1}, \qquad \varepsilon_{g,t} \sim N(0, \sigma_g^2).$$

(This is a good approximation to US post-war data.[15]) Combining these two equations gives that:

$$r_t = -\log\beta + \frac{1-\rho_g}{\varsigma} g - \frac{1}{2}\left(\frac{\sigma_g}{\varsigma}\right)^2 + \frac{\rho_g}{\varsigma} g_t + \frac{\theta_g}{\varsigma}\varepsilon_{g,t},$$

implying that a (roughly) $\frac{\rho_g}{\varsigma}$ response to consumption growth can substitute for a (roughly) unit response to real rates. Of course, output (growth, level or gap) is in turn highly correlated with consumption growth, so output (growth, level or gap) may also substitute for real rates. For example, in the Smets & Wouters (2007) model of the US economy, the monetary rule is of the form $i_t = \phi_\pi \pi_t + z_t + \zeta_t$, where z_t is a linear combination of other endogenous variables and ζ_t is the monetary shock. At the estimated posterior mode, the correlation between z_t and the real interest rate is 0.63, with both variables having standard deviation of 0.46%. Thus, the Smets & Wouters (2007) estimates imply that the Fed is already about two thirds of the way to using a robust real rate rule.

1.3 Implementing arbitrary inflation dynamics

Real rate rules can determinately implement any path for inflation, no matter the rest of the model. Let π_t^* be an exogenous stochastic process, perhaps a function of the economy's other shocks,[16] and consider the rule:

$$i_t = r_t + \mathbb{E}_t \pi_{t+1}^* + \phi(\pi_t - \pi_t^*). \tag{7}$$

From the Fisher equation (1), this implies:

---

[15] Estimating on US data from 1947Q1 to 2021Q4 (BEA series: A794RX) with T-distributed shocks gives $\rho_g = 0.69$, $\theta_g = -0.50$ with p-values both below 10^{-5}. Estimating with Gaussian shocks on less volatile sub-periods gives similar results.

[16] Ireland (2007) also allows the central bank's inflation target to respond to other structural shocks.

$$\mathbb{E}_t(\pi_{t+1} - \pi^*_{t+1}) = \phi(\pi_t - \pi^*_t).$$

Again with $\phi > 1$, there is a unique solution, now with $\pi_t = \pi^*_t$. I.e., at all periods of time, and in all states of the world, realised inflation is equal to π^*_t. Effectively, the central bank is able to choose an arbitrary path for inflation as the unique, determinate equilibrium outcome.

The only constraint is that the targeted path for inflation cannot be a function of endogenous variables. However, this is not much of a limitation, since in stationary equilibrium, endogenous variables must have a representation as a function of the infinite history of the economy's shocks. This means that by choosing π^*_t appropriately, rules in the form of (6) can mimic the outcomes of any other monetary policy regime.[17]

For example, suppose that the central bank were to set interest rates in a different (though time invariant) way, for example by using another rule, or by adopting optimal policy under either commitment or discretion, given some objective. For simplicity, suppose further that the economy's equilibrium conditions are linear, e.g., because we are working under a first order approximation. Let $(\varepsilon_{1,t}, \dots, \varepsilon_{N,t})_{t\in\mathbb{Z}}$ be the set of structural shocks in the economy,[18] all of which are assumed mean zero and independent both of each other, and over time. Finally, assume that the central bank's behaviour produces stationary inflation, $\tilde{\pi}_t$, with the $\widetilde{\ }$ denoting that this is inflation under the alternative monetary regime. Then, by linearity and stationarity, there must exist a constant $\tilde{\pi}^*$ and coefficients $(\theta_{1,k}, \dots, \theta_{N,k})_{k\in\mathbb{N}}$ such that:

$$\tilde{\pi}_t = \tilde{\pi}^* + \sum_{k=0}^{\infty} \sum_{n=1}^{N} \theta_{n,k}\varepsilon_{n,t-k},$$

with $\sum_{k=0}^{\infty} \theta^2_{n,k} < \infty$ for $n = 1, \dots, N$. So, if the central bank sets:

---

[17] Other papers have examined the implementation of optimal policy in specific models using instrument rate rules (see e.g. Svensson & Woodford 2003; Dotsey & Hornstein 2006; Evans & Honkapohja 2006; Evans & McGough 2010). However, the various prior proposals do not enable the implementation of a certain inflation path robustly across models.

[18] This may include sunspot shocks if they are added following Farmer, Khramov & Nicolò (2015).

$$\pi_t^* = \tilde{\pi}^* + \sum_{k=0}^{\infty} \sum_{n=1}^{N} \theta_{n,k} \varepsilon_{n,t-k},$$

(exogenous!) and uses the rule (6), then for all t and in all states of the world, $\pi_t = \pi_t^* = \tilde{\pi}_t$. Moreover, this implies in turn that all the endogenous variables in the two economies must be identical in all periods and in all states of the world.[19]

The only slight difficulty with this approach is that the central bank may struggle to observe structural shocks. The central bank can certainly observe linear combinations of structural shocks, via estimating a VAR with sufficiently many lags. For variables that are plausibly contemporaneously exogenous, such as commodity prices for a small(ish) economy, this is already sufficient to recover the corresponding structural shock. To infer other shocks, the central bank needs to know more about the structure of the economy. However, we do not need to assume any more than is standard in rational expectations models. Forming rational expectations requires you to know the structure of the economy; if you know this structure, then you know the mapping from the reduced form shocks estimated by a VAR to the model's structural shocks.[20]

2 Practical implementation of real rate rules

Until recently, central banks concentrated their monetary interventions in overnight debt markets. However, with the rise of quantitative easing, many central banks have been purchasing substantial quantities of longer maturity sovereign debt. There is no reason then that central banks could not conduct open market operations to fix the interest rate on longer maturity bonds. This is convenient as in most countries, inflation protected securities are only issued a few times per year, and at long maturities, e.g., five years. As a result, markets in shorter maturity inflation protected securities may be illiquid or even unavailable, and it can be difficult to reconstruct the

[19] Proven in Appendix A.5.

[20] This mapping may not be unique valued if there are more shocks than observables. However, since we expect a relatively small number of shocks to explain the bulk of business cycle variance, this is unlikely to be problematic in practice.

short end of the real yield curve. Inflation indexation lags further complicate the use of short maturity inflation protected securities (see e.g. Gürkaynak, Sack & Wright (2010)). For example, 3-month maturity US treasury inflation protected securities (TIPS) have a period t realized yield of $r_{t-1} + \pi_{t-1}$, not $r_{t-1} + \pi_t$ as one would hope, where time is measured in quarters.

In practice, the central bank's trading desk would be tasked with maintaining a particular level of the gap between nominal and real rates according to the market for bonds of a certain maturity. For the rest of this section, we shall assume five-year bonds are used, since five-year TIPS are the shortest maturity issued in the US.

So, let i_t be the nominal yield per-period on a five-year sovereign bond at t, and r_t be the real yield per-period on a five-year inflation protected bond from the same issuer. As ever, t indexes time. The units of time do not need to coincide with the maturity of the bond. In particular, t may be measured in months, quarters or years, in which case i_t is the nominal yield per-month, per-quarter or per-year, respectively. Let T be the number of periods in five years. For example, T may be 60 if periods are months.

We also allow for the possibility that inflation is not observed contemporaneously. For example, US CPI is observed with a one-month lag. To capture this, while keeping to the convention that $\mathbb{E}_t v_t = v_t$ for all t-dated endogenous variables v_t, we assume that market participants and the central bank use the $t - L$ information set in period t (i.e. they know the values of all $t - L, t - L - 1, \ldots$ dated variables), for some $L \geq 0$. Thus, since the central bank does not know π_t at t, we instead assume that they respond to deviations of π_{t-L} from target, rather than π_t.

We allow for a shock in the Fisher equation to capture inflation risk premia, liquidity premia, asymmetric term premia and even further departures from full information rational expectations amongst market participants. Since only $t - L$ dated variables are known in period t, we denote the period t value of this shock by ν_{t-L}. I.e., risk premia (etc.) will be determined L periods in advance, though market participants and the central bank will not act on this, since they use L period old data. Given this, the Fisher equation coming from arbitrage between nominal and real bonds then states that:

$$i_t - r_t = \nu_{t-L} + \mathbb{E}_{t-L} \frac{1}{T} \sum_{k=1}^{T} \pi_{t+k},$$

where ν_{t-L} is the aforementioned shock to risk premia (etc.). We only require that ν_t is a stationary process.

TODO Allow for indexation lags in TIPS

Slightly generalizing our previous rule (6), we suppose that the central bank intervenes in five-year nominal bond markets to ensure that it is always the case that:

$$i_t - r_t = \bar{\nu}_{t-L} + \mathbb{E}_{t-L} \frac{1}{T} \sum_{k=1}^{T} \pi_{t+k}^{*} + \phi(\pi_{t-L} - \pi_{t-L}^{*}),$$

where $\bar{\nu}_{t-L}$ is the central bank's period t belief about the level of ν_{t-L}.

We have deliberately not added any interest rate smoothing. While such smoothing is often believed to be a relevant feature of real-world central bank behaviour, in our context it adds nothing. Smooth paths for interest rates may be produced from a smooth target path for π_t^*.[21]

Also note that while under conventional monetary policy, targeted nominal interest rates are (approximately) constant between monetary policy committee meetings, this may not be the case here. The rule effectively specifies a period t level for $i_t - r_t$, not for i_t. The level of r_t may fluctuate (perhaps in part due to unexpected changes in i_t), so the central bank's trading desk could have to continuously tweak the level of i_t to hold $i_t - r_t$ at its desired level. While this represents a departure from previous operating procedure, there is no reason why holding $i_t - r_t$ approximately constant should be any harder than holding i_t approximately constant. This is thanks to real-time observability of r_t via inflation protected bonds. The central bank could also

[21] In situations in which the dynamics of π_t^* are constrained, then adding smoothing may help match real-world dynamics. In this case, the independence of inflation from the rest of the economy can be preserved if rather than i_{t-1} appearing on the right hand side of the monetary rule, instead there is $i_{t-1} - r_{t-1}$.

directly control $i_t - r_t$ by promising to freely exchange \$1 face value of real debt for $\$(1 + i_t - r_t)$ face value of nominal debt, as suggested by Cochrane (2017; 2018). Alternatively, the central bank could buy or sell a long-short portfolio containing \$1 face value of nominal debt, and −\$1 face value of real debt to hold the portfolio's price fixed at $\$(i_t - r_t)$.[22]

Thus, the monetary rule implies that the dynamics of inflation are governed by the single equation:

$$\mathbb{E}_{t-L} \frac{1}{T} \sum_{k=1}^{T} (\pi_{t+k} - \pi^*_{t+k}) = (\bar{\nu}_{t-L} - \nu_{t-L}) + \phi(\pi_{t-L} - \pi^*_{t-L}),$$

i.e.:

$$\mathbb{E}_t \frac{1}{T} \sum_{k=1}^{T} (\pi_{t+k+L} - \pi^*_{t+k+L}) = (\bar{\nu}_t - \nu_t) + \phi(\pi_t - \pi^*_t).$$

As ever, with $\phi > 1$, there is a unique solution.[23] In the special case in which the central bank observes ν_t (i.e. risk premia etc.) so $\nu_t = \bar{\nu}_t$, then $\pi_t = \pi^*_t$, as before. In the general case, as long as $\bar{\nu}_t - \nu_t$ is stationary, the solution takes the form:[24]

$$\pi_t = \pi^*_t + \mathbb{E}_t \sum_{j=0}^{\infty} A_j (\bar{\nu}_{t+j} - \nu_{t+j}),$$

where $A_0 \coloneqq -\frac{1}{\phi}$, $A_j \coloneqq 0$ for $j \in \{1, \dots, L\}$, and $A_j \coloneqq \frac{1}{\phi T} \sum_{k=\max\{0, j-L-T\}}^{j-L-1} A_k$ for all $j > L$, implying $A_{L+1} = -\frac{1}{T\phi^2}$ and $A_j = \mathrm{O}\left(\phi^{-\frac{j}{T+L}}\right)$ as $j \to \infty$.[25] Thus, with ϕ large, even if the central bank imperfectly tracks the risk (etc.) premium ν_t, it will still be the case that $\pi_t \approx \pi^*_t$ in all periods. I.e., even in the presence of unobservable risk premia, the central bank can still determinately implement an arbitrary path for inflation. The presence of information lags makes no fundamental difference to this. While

[22] The author thanks Peter Ireland for this suggestion.

[23] We do not have the indeterminacy issues for rules setting long-rates that were noted by McGough, Rudebusch & Williams (2005), due to the presence of the real rate in our rule.

[24] Ireland (2015) finds a role for risk premia in explaining US inflation fluctuations, so risk premia appearing in the solution for inflation should not be too surprising.

[25] Guess $A_j \propto B^j$. Then (for large j): $B^j = \frac{1}{\phi T} \sum_{k=j-L-T}^{j-L-1} B^k = \frac{1}{\phi T} \frac{B^{j-L-T} - B^{j-L}}{1-B}$, so $\phi T B^{T+L} = \frac{1-B^T}{1-B} \in [1, T]$, implying $0 \le B \le \phi^{-\frac{1}{T+L}}$.

information lags may slow down the convergence of A_j to 0 as $j \to \infty$, increasing the variance of $\pi_t - \pi_t^*$, still for a large enough ϕ, inflation will be very close to its target.

TODO Irregular meetings.

TODO Geometric bonds and perpetuities.

TODO Inflation swaps (more liquid, no deflation protection so simpler pricing)

3 Real rate rules in a model of the redistribution channel of monetary policy

Sticky prices or wages are not the only channel by which monetary policy may come to have real effects. With households' wealth or debt held in nominal bonds, unexpected inflation leads to redistribution from savers to borrowers, as shown empirically by Doepke & Schneider (2006). The wealth effect will then lead savers to work more and borrowers to work less. If saver households labour supply is more responsive to income than that of borrowers, then total labour supply can increase, producing a positive relationship between output and inflation, even with flexible prices. This mechanism is at work in many monetary heterogeneous agent models (see e.g. Kaplan, Moll & Violante (2018) and Auclert (2019)), though it may be obscured by sticky prices. In this section we present the simplest possible model of this redistribution channel and use it to further examine the behaviour of real rate rules.

Examining this model is helpful for two main reasons. Firstly, it demonstrates that there is nothing about the performance of real rate rules that is specific to variants of the New Keynesian model. With the increased prominence of heterogenous agent models in the monetary literature, it is particularly important to check robustness to the redistribution channel.

Secondly, our model of the redistribution channel is so simple that it has an exact analytic solution even in the presence of stochastic volatility. So far, we have examined real rate rules either in models in which the linearized Fisher equation ($i_t = r_t + \mathbb{E}_t\pi_{t+1}$) holds exactly, or in models in which it contains an exogenous risk premium term. However, risk premia are potentially endogenous. As such, they may affect determinacy conditions, or otherwise change the performance of real rate rules. By including stochastic volatility, our model will generate an endogenous time varying risk premium, allowing us to assess whether real rate rules can retain their good properties in the presence of such endogenous risk premia in the Fisher equation.

3.1 The model

We now describe the model. The final good is produced under perfect competition using the linear technology $Y_t = Z_t l_t$, where Z_t is an exogenous productivity process, and l_t is aggregate labour supply. Labour is paid a real wage W_t. The final good producer's first order condition implies $W_t = Z_t$.

We define productivity growth $G_t := \frac{Z_t}{Z_{t-1}}$ and assume that G_t evolves according to the law of motion:

$$\gamma_t := \log\left(\frac{G_t}{G}\right) = \rho_\gamma \gamma_{t-1} + (1-\rho_\gamma) s_{\gamma,t-1}\varepsilon_{\gamma,t}.$$

Here $\varepsilon_{\gamma,t}$ is a shock drawn from a time invariant symmetric distribution with $\mathbb{E}\varepsilon_{\gamma,t} = 0$ and $\sigma_\gamma^2 := \mathbb{E}\varepsilon_{\gamma,t}^2 > 0$. We further assume that for all t, $\varepsilon_{\gamma,t} \in [-1,1]$. Bounded shocks will help the model's tractability. The distribution of $\varepsilon_{\gamma,t}$ is otherwise unrestricted.

$s_{\gamma,t}$ is an exogenous process determining the volatility of γ_t. We will define the law of motion for $s_{\gamma,t}$ in terms of the cumulant generating function of $\varepsilon_{\gamma,t}$. This is the function $m_\gamma \colon \mathbb{R} \to [0,\infty)$ given by:

$$m_\gamma(z) = \log \mathbb{E}\exp(z\varepsilon_{\gamma,t}),$$

for all $z \in \mathbb{R}$. Our assumptions imply that m_γ is convex, that $m_\gamma(-z) = m_\gamma(z)$ for all z, and that $m_\gamma(z) = \frac{1}{2}\sigma_\gamma^2 z^2 + O(z^4)$ as $z \to 0$. Since $m_\gamma(z)$ is strictly increasing on $[0,\infty)$, it possesses an inverse $m_\gamma^{-1}\colon [0,\infty) \to [0,\infty)$. Using this, we define the following law of motion for volatility $s_{\gamma,t}$:

$$s_{\gamma,t} = \frac{1}{1-\rho_\gamma} m_\gamma^{-1}\Big(\varrho_\gamma m_\gamma\Big((1-\rho_\gamma)s_{\gamma,t-1}\Big) + (1-\varrho_\gamma) m_\gamma\Big((1-\rho_\gamma)\bar{s}_\gamma\Big)\eta_{\gamma,t}\Big),$$

where the volatility innovation $\eta_{\gamma,t}$ is drawn from another time invariant distribution with $\eta_{\gamma,t} \in [0,1]$ for all t. To see why this process is less weird than it may at first seem, define $v_{\gamma,t} \coloneqq m_\gamma\Big((1-\rho_\gamma)s_{\gamma,t}\Big)$ and $\bar{v}_\gamma \coloneqq m_\gamma\Big((1-\rho_\gamma)\bar{s}_\gamma\Big)$. Then the previous law of motion for $s_{\gamma,t}$ implies that:

$$v_{\gamma,t} = \varrho_\gamma v_{\gamma,t-1} + (1-\varrho_\gamma)\bar{v}_\gamma \eta_{\gamma,t}.$$

By the previous approximation result for m_γ, $v_{\gamma,t} = \left[\frac{1}{2}(1-\rho_\gamma)^2\sigma_\gamma^2\right]s_{\gamma,t}^2 + O(\bar{s}_\gamma^4)$ and $\bar{v}_\gamma = \left[\frac{1}{2}(1-\rho_\gamma)^2\sigma_\gamma^2\right]\bar{s}_\gamma^2 + O(\bar{s}_\gamma^4)$ as $\bar{s}_\gamma \to 0$. So, we have that:

$$s_{\gamma,t}^2 = \varrho_\gamma s_{\gamma,t-1}^2 + (1-\varrho_\gamma)\bar{s}_\gamma^2 \eta_{\gamma,t} + O(\bar{s}_\gamma^4)$$

as $\bar{s}_\gamma \to 0$. I.e., the squared volatility roughly follows an AR(1) process. Our chosen law of motion for $s_{\gamma,t}$ will permit an exact solution, so we will not use this approximation in the below. Finally note that by construction, $s_{\gamma,t} \le \bar{s}_\gamma$, so $\gamma_t \in [-\bar{s}_\gamma, \bar{s}_\gamma]$ for all t.

There is a unit mass of households, indexed by i. Households only differ in their discount factor, $\beta_i \in (0,1)$. Each household has a unit mass of members, who may each supply either 0 or 1 units of labour (as in Hansen (1985)). Each working member incurs disutility of $\chi > 0$. The household head determines what proportion of its members work and allocates the same consumption to all household members. As a result, in period t, the head of household i maximizes:

$$\mathbb{E}_t \sum_{k=0}^{\infty} \beta_i^k [\log C_{i,t+k} - \chi l_{i,t+k}],$$

where $C_{i,t} \ge 0$ is their real consumption and $l_{i,t} \in [0,1]$ is their labour supply, i.e. the proportion of household members who work. The maximization is subject to the budget constraint:

$$C_{i,t} + A_{i,t} + B_{i,t} = W_t l_{i,t} + A_{i,t-1}R_{t-1} + B_{i,t-1}\frac{I_{t-1}}{\Pi_t},$$

where $A_{i,t}$ is the real quantity of real bonds they purchase in period t, which have real return R_t in period $t+1$, and $B_{i,t}$ is the real quantity of nominal bonds they purchase in period t, which have real return $\frac{I_t}{\Pi_{t+1}}$ in period $t+1$. Here, R_t is the gross real interest rate from t to $t+1$, I_t is the gross nominal interest rate from t to $t+1$, and Π_t is the gross inflation rate from $t-1$ to t. As usual, we define $\pi_t := \log \Pi_t$.

Households face both real and nominal borrowing constraints. They cannot borrow at all in real bonds, so $A_{i,t} \geq 0$. This is in line with the rarity of inflation protected debt contracts in reality. With nominal bonds, they face the following debt to income constraint:

$$B_{i,t} \geq -\psi W_t l_{i,t},$$

where $\psi > 0$. The Bank of England imposed a borrowing constraint of this form in 2014.[26] Constraints on expected debt payments, i.e., with $\left(\mathbb{E}_t \frac{I_t}{\Pi_{t+1}} - 1\right) B_{i,t}$ on the left hand side, are almost equivalent to this constraint, as $\mathbb{E}_t \frac{I_t}{\Pi_{t+1}}$ will not vary much in equilibrium. Borrowing constraints of roughly this form are common in the US.

Since Z_t is non-stationary, it is convenient to rewrite the household's problem in terms of detrended variables. We define $c_{i,t} := \frac{C_{i,t}}{Z_t}$, $a_{i,t} := \frac{A_{i,t}}{Z_t}$, $b_{i,t} := \frac{B_{i,t}}{Z_t}$, $w_t := \frac{W_t}{Z_t} = 1$ and $y_t := \frac{Y_t}{Z_t} = l_t$. Then the household's budget constraint may be rewritten as:

$$c_{i,t} + a_{i,t} + b_{i,t} = l_{i,t} + a_{i,t-1} \frac{R_{t-1}}{G_t} + b_{i,t-1} \frac{I_{t-1}}{\Pi_t G_t}.$$

The other constraints are that $l_{i,t} \leq 1$, $a_{i,t} \geq 0$ and $b_{i,t} \geq -\psi l_{i,t}$. Let $\lambda_{i,t}$, $\mu_{i,t}$ and $\nu_{i,t}$, respectively, be the Lagrange multipliers on these last three constraints. Then the household's first order conditions imply:

$$\frac{1}{c_{i,t}} + \psi \nu_{i,t} = \chi + \lambda_{i,t}, \qquad 0 = \min\{\lambda_{i,t}, 1 - l_{i,t}\},$$

$$\mu_{i,t} + \beta_i R_t \mathbb{E}_t \frac{1}{G_{t+1} c_{i,t+1}} = \frac{1}{c_{i,t}}, \qquad \nu_{i,t} + \beta_i I_t \mathbb{E}_t \frac{1}{\Pi_{t+1} G_{t+1} c_{i,t+1}} = \frac{1}{c_{i,t}},$$

$$0 = \min\{\mu_{i,t}, a_{i,t}\}, \qquad 0 = \min\{\nu_{i,t}, b_{i,t} + \psi l_{i,t}\}.$$

[26] https://www.reuters.com/article/idUKKBN0F10UE20140626

To keep the model as simple as possible, we suppose there are in fact only two types of households (as in Galí, López-Salido & Vallés (2007) and Bilbiie (2008)). A fraction $1-\theta$ of households are type "B", with $\beta_i = \beta_B$. A fraction θ of households are type "S", with $\beta_i = \beta_S$. We will suppose that $\beta_S > \beta_B$, so S-type households will be savers in equilibrium, while B-type households will be borrowers. In equilibrium, all B-type households will make the same decisions, and so too will all S-type households. Thus, we drop "i" subscripts and replace them with "B" or "S" subscripts to represent the decisions of B-type and S-type households respectively. Market clearing then implies that:

$$(1-\theta)c_{B,t} + \theta c_{S,t} = y_t = l_t = (1-\theta)l_{B,t} + \theta l_{S,t},$$

$$0 = (1-\theta)a_{B,t} + \theta a_{S,t}, \qquad 0 = (1-\theta)b_{B,t} + \theta b_{S,t}.$$

We look for an equilibrium in which in all periods, B-type households supply as much labour as they can ($l_{B,t} = 1$), have zero real bonds ($a_{B,t} = 0$) and borrow as much as they can in nominal bond markets ($b_{B,t} = -\psi$). If B-type households have zero real bonds, then by market clearing, S-type households must also have zero real bonds ($a_{S,t} = 0$). Similarly, if B-type households have nominal debt, then S-type households must have nominal assets ($b_{S,t} > 0$). In fact, market clearing implies that for all t:

$$b_{S,t} = \psi \frac{1-\theta}{\theta}.$$

As S-type households are richer than B-type households, they will have a lower incentive to supply labour. We look for an equilibrium in which S-type households are not at the labour supply bound (i.e., $\lambda_{S,t} = 0$ for all t). With B-type households at the labour supply limit, but S-type households below it, monetary expansions will lead S-type households to work more (due to inflation eroding real wealth) but will not change the labour supply of B-type households. The same core mechanism would work without indivisible labour as long as labour supply becomes less elastic the higher it is, for example due to a labour supply disutility of the form $\log(1 - l_{i,t})$.[27] Our chosen preferences give greater tractability.

Substituting our equilibrium restrictions into the market clearing conditions and the household's first order constraints and budget constraint gives that:

$$
\begin{aligned}
c_{B,t} &= 1 - \psi\left(\frac{I_{t-1}}{\Pi_t G_t} - 1\right), \qquad c_{S,t} = \frac{1}{\chi}, \\
l_{S,t} &= \frac{1}{\chi} - \psi\frac{1-\theta}{\theta}\left(\frac{I_{t-1}}{\Pi_t G_t} - 1\right), \\
y_t = l_t &= (1-\theta) + \theta\frac{1}{\chi} - \psi(1-\theta)\left(\frac{I_{t-1}}{\Pi_t G_t} - 1\right), \qquad (8) \\
\beta_S R_t \mathbb{E}_t \frac{1}{G_{t+1}} &= 1, \qquad \beta_S I_t \mathbb{E}_t \frac{1}{\Pi_{t+1} G_{t+1}} = 1.
\end{aligned}
$$

The consumption of B-type households varies negatively with $\frac{I_{t-1}}{\Pi_t G_t}$. When inflation is unexpectedly high, B-type households have less real debt to repay, and so they wish to consume more. To support the higher consumption of B-type households when inflation is high, S-type households work more at these times. The consumption of S-type households is proportional to productivity, Z_t. There is still a non-trivial stochastic discount factor of $\frac{\beta_S}{G_{t+1}}$ as Z_t is stochastic. S-type households are able to insure themselves so well due to their highly elastic labour supply. In richer models, wealthy households are still well insured against inflationary shocks, due to their real asset income dominating that from nominal assets or labour.

Since unexpected inflation leads B-type households to consume more (due to lower real debt) and S-type households to work more (due to lower real wealth), the model produces a Phillips curve type relationship between output and inflation. Substituting the nominal bond first order condition into the output equation (8) gives:

$$
y_t = (1-\theta) + \theta\frac{1}{\chi} - \psi(1-\theta)\left(\frac{1}{\beta_S}\frac{1}{\Pi_t G_t}\left[\mathbb{E}_{t-1}\frac{1}{\Pi_t G_t}\right]^{-1} - 1\right).
$$

So, to a first order approximation:

$$
y_t \approx y + \frac{\psi(1-\theta)}{\beta_S}[\pi_t - \mathbb{E}_{t-1}\pi_t + \gamma_t - \mathbb{E}_{t-1}\gamma_t].
$$

[27] With this utility, the Frisch elasticity of labour supply is $\frac{1-l_{i,t}}{l_{i,t}}$, which falls from ∞ to 0 as $l_{i,t}$ increases from 0 to 1.

This is a New Classical type Phillips curve.

TODO Show plot of this Phillips curve

For the equilibrium to make sense, we need $l_{S,t} \in [0,1]$, $c_{B,t} \geq 0$ and $\lambda_{B,t} \geq 0$ for all t. $l_{S,t} \in [0,1]$ requires:

$$\frac{\chi\psi(1-\theta)}{\chi\psi(1-\theta)+\theta} \leq \frac{\Pi_t G_t}{I_{t-1}} \leq \frac{\chi\psi(1-\theta)}{\chi\psi(1-\theta)-(\chi-1)\theta}.$$

These two inequalities are sufficient for $c_{B,t} \geq 0$ if $\frac{\theta}{1-\theta} \leq \chi$, and they are sufficient for $\lambda_{B,t} \geq 0$ if $\chi \leq 1 + (1-\theta)\psi$ and β_B is sufficiently small.[28] For the inequalities to hold for all t, with probability one, we need $\frac{\Pi_t G_t}{I_{t-1}}$ to be bounded. For this, it is sufficient that $\bar{s}_\gamma$ and $\bar{s}_\zeta$ are small enough, where $\bar{s}_\zeta$ is the analogous parameter to $\bar{s}_\gamma$ for the monetary policy shock process, to be defined.

TODO complete write up

4 Welfare

In Section 1, we established that a rule of our form could exactly mimic any other time invariant policy, if responses to structural shocks and their lags are allowed. Thus, rules of our form can mimic unconditionally optimal policy, optimal commitment policy from a timeless perspective, or optimal discretionary policy. Hence, rules of our form can achieve high welfare.

We begin this section by looking at unconditionally optimal time-invariant policy using our rules, in a simple model. We then go on to analyse the performance of our rules if further restrictions are placed upon them, such as only permitting the central bank to respond to current or sufficiently recent shocks. We show that optimal policy in estimated models of the US economy comes close to stabilizing inflation, with optimal inflation dynamics describable by an ARMA process with few MA terms.

[28] β_B does not otherwise enter the equilibrium conditions, so we are free to see $\beta_B \approx 0$.

Any welfare analysis requires us to specify the rest of the model, as welfare is generally a function of output's variability, not just that of inflation. Thus, as a first example suppose that inflation and output are linked by the standard Phillips curve:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + \kappa \omega_t,$$

where x_t is the output gap, and ω_t is a mark-up shock, which is assumed IID with mean zero. Additionally, suppose that the policy maker wants to minimise the unconditional expectation of a quadratic loss function in inflation and the output gap. I.e. the period t policy maker minimises:

$$(1-\beta)\mathbb{E}\sum_{k=0}^{\infty} \beta^k \left(\pi_{t+k}^2 + \lambda x_{t+k}^2\right),$$

for some $\lambda > 0$ and $\beta \in (0,1)$.

We suppose that the policy maker is constrained to choose a time-invariant (i.e. stationary) policy, thus the objective simplifies to:[29]

$$\mathbb{E}\left(\pi_t^2 + \lambda x_t^2\right).$$

As the policy maker only cares about inflation and output gaps, with the former being effectively under their control, and the latter only determined by inflation and mark-up shocks, the optimal policy must have the form:

$$\pi_t = \kappa \sum_{k=0}^{\infty} \theta_k \omega_{t-k},$$

for some $\theta_0, \theta_1, \dots$ to be determined. We have already shown that such a policy may be determinately implemented via a rule of the form of (6).

Substituting this policy into the Phillips curve then gives:

$$\sum_{k=0}^{\infty} \theta_k \omega_{t-k} = \beta \sum_{k=0}^{\infty} \theta_{k+1} \omega_{t-k} + x_t + \omega_t,$$

so:

$$x_t = \sum_{k=0}^{\infty} \left(\theta_k - \beta\theta_{k+1} - \mathbb{1}[k=0]\right)\omega_{t-k}.$$

Hence, the policy maker's objective is to choose $\theta_0, \theta_1, \dots$ to minimise:

[29] See e.g. Damjanovic, Damjanovic & Nolan (2008).

$$\mathbb{E}\left(\pi_t^2+\lambda x_t^2\right)=\mathbb{E}\left[\omega_t^2\right]\sum_{k=0}^{\infty}\left[\kappa^2\theta_k^2+\lambda(\theta_k-\beta\theta_{k+1}-\mathbb{1}[k=0])^2\right].$$

The first order conditions then give:[30]

$$\begin{aligned}
&\theta_0+\frac{\lambda}{\kappa^2}(\theta_0-\beta\theta_1-1)=0,\\
&\theta_1+\frac{\lambda}{\kappa^2}(\theta_1-\beta\theta_2)-\beta\frac{\lambda}{\kappa^2}(\theta_0-\beta\theta_1-1)=0,\\
\forall k>1,\qquad &\theta_k+\frac{\lambda}{\kappa^2}(\theta_k-\beta\theta_{k+1})-\beta\frac{\lambda}{\kappa^2}(\theta_{k-1}-\beta\theta_k)=0.
\end{aligned}$$

Unsurprisingly, this agrees with the unconditionally optimal solution given in the prior literature (e.g. Damjanovic, Damjanovic & Nolan (2008)), which satisfies:

$$\pi_t+\frac{\lambda}{\kappa}(x_t-\beta x_{t-1})=0,$$

i.e.:

$$\kappa\sum_{k=0}^{\infty}\theta_k\omega_{t-k}+\frac{\lambda}{\kappa}\left[\sum_{k=0}^{\infty}(\theta_k-\beta\theta_{k+1}-\mathbb{1}[k=0])\omega_{t-k}\right.\\
\left.-\beta\sum_{k=1}^{\infty}(\theta_{k-1}-\beta\theta_k-\mathbb{1}[k-1=0])\omega_{t-k}\right]=0.$$

To see the equivalence, note that from matching coefficients, this equation holds if and only if the above first order conditions hold. We will present a convenient representation of the solution to these equations below.

Additionally, note that as $\frac{\lambda}{\kappa^2}\to 0$, $\theta_k\to 0$ for all $k\in\mathbb{N}$. In other words, if the central bank does not care about the output gap, then they optimally choose to have constant inflation, i.e., to follow the rule from equation (2). The central bank also chooses constant inflation if the Phillips curve is vertical (i.e. $\kappa=\pm\infty$). In this case, neither inflation nor mark-up shocks have any impact on the output gap.

The first order conditions derived above also enable us to easily solve for optimal unconditional policy under limited memory. For example, if the central bank does not "remember" $\omega_{t-1},\omega_{t-2},\ldots$, so uses a rule that is only a function of ω_t at t, then the optimal θ_0 will satisfy the above first order conditions with $\theta_1=\theta_2=\cdots=0$. This means:

[30] See Appendix A.6 for the solution of these conditions.

$$\theta_0 + \frac{\lambda}{\kappa^2}(\theta_0 - 1) = 0,$$

so $\theta_0 = \frac{\lambda}{\lambda+\kappa^2}$. It turns out that this exactly coincides with the solution under discretion.[31]

If the central bank can "remember" ω_{t-1}, so π_t is an MA(1), then the optimal solution will have:

$$\theta_0 + \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$
$$\theta_1 + \frac{\lambda}{\kappa^2}\theta_1 - \beta\frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0.$$

The solution has $\theta_0 \geq 0$ and $\theta_1 \leq 0$. Thus, the shock increases π_t while reducing $\mathbb{E}_t\pi_{t+1}$, thus dampening the required movement in x_t, from the Phillips curve. We will see that this is already enough to come close to the fully optimal policy.

Going one step further, if the central bank can also "remember" π_{t-1}, then they can choose interest rates to ensure π_t follows the ARMA(1,1) process:

$$\pi_t = \rho\pi_{t-1} + \kappa\theta_0\omega_t + \kappa\theta_1\omega_{t-1},$$

for some ρ, θ_0, θ_1 to be determined.[32] Since US inflation appears to be well approximated by an ARMA(1,1) (Stock & Watson 2009), this may be a reasonable model of Fed behaviour. This ARMA(1,1) process has the MA(∞) representation:

$$\pi_t = \kappa\theta_0\sum_{k=0}^{\infty}\rho^k\omega_{t-k} + \kappa\theta_1\sum_{k=0}^{\infty}\rho^k\omega_{t-1-k} = \kappa\theta_0\omega_t + \kappa(\rho\theta_0 + \theta_1)\sum_{k=1}^{\infty}\rho^{k-1}\omega_{t-k}. \tag{9}$$

Substituting this policy into the Phillips curve gives:

$$\theta_0\omega_t + (\rho\theta_0 + \theta_1)\sum_{k=1}^{\infty}\rho^{k-1}\omega_{t-k} = \beta(\rho\theta_0 + \theta_1)\omega_t + \beta(\rho\theta_0 + \theta_1)\sum_{k=1}^{\infty}\rho^k\omega_{t-k} + x_t + \omega_t,$$

so:

$$x_t = [(1-\beta\rho)\theta_0 - \beta\theta_1 - 1]\omega_t + (1-\beta\rho)(\rho\theta_0 + \theta_1)\sum_{k=1}^{\infty}\rho^{k-1}\omega_{t-k}.$$

Hence, the policy maker's objective is to choose ρ, θ_0, θ_1 to minimise:

[31] See Appendix A.7.

[32] The targeted inflation can respond to lagged targeted inflation without changing the determinacy properties of realised inflation (always equal to targeted inflation in equilibrium). Targeted inflation cannot respond to other endogenous variables without potentially changing these determinacy properties.

$$\mathbb{E}(\pi_t^2 + \lambda x_t^2) = \mathbb{E}[\omega_t^2]\Bigg[\kappa^2\theta_0^2 + \lambda[(1-\beta\rho)\theta_0 - \beta\theta_1 - 1]^2 + [\kappa^2(\rho\theta_0 + \theta_1)^2 + \lambda(1-\beta\rho)^2(\rho\theta_0+\theta_1)^2]\frac{1}{1-\rho^2}\Bigg].$$

Tedious algebra gives that the first order conditions have solution:[33]

$$\rho = \frac{\kappa^2 + (1+\beta^2)\lambda - \sqrt{(\kappa^2 + (1-\beta)^2\lambda)(\kappa^2 + (1+\beta)^2\lambda)}}{2\beta\lambda}, \qquad \theta_0 = \frac{\rho}{\beta}, \qquad \theta_1 = -\rho.$$

As $\lambda \to 0$, or $\kappa \to \infty$, $\rho \to 0$. As $\lambda \to \infty$, or $\kappa \to 0$, $\rho \to \beta$. Since there is no other solution for κ to the equation $\rho = \beta$ than $\kappa = 0$, we must have $\rho \leq \beta$, so $\rho\theta_0 + \theta_1 \leq 0$, meaning that the response of inflation to a positive mark-up shock is again negative after the first period. Since we have one extra degree of freedom, this must attain even higher welfare than the MA(1) solution. In fact, it attains the unconditionally optimal solution. Examination of the unconditionally optimal solution from Appendix A.6 reveals that it has the same form as equation (8), thus by a revealed preference argument, the two solutions must coincide. (For example, the solution for ρ agrees with the geometric decay rate of the MA coefficients at lags beyond the first of the fully optimal solution we found in Appendix A.6.)

Hence, in a world in which the only inefficient shocks are IID cost-push shocks, the central bank can attain the unconditionally optimal welfare by ensuring inflation follows an appropriate ARMA(1,1) process. This process will have an MA coefficient equal to $-\beta \approx -0.99$, and as long as the central bank cares about output stabilisation, it will have a high degree of persistence. This is very close to the IMA(1,1) processes estimated by Dotsey, Fujita & Stark (2018) for the post-1984 period.

[33] There is an additional solution to the first order condition with $\rho = \frac{\kappa^2+(1+\beta^2)\lambda+\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1+\beta)^2\lambda)}}{2\beta\lambda}$, but this is outside of the unit circle as: $\frac{\kappa^2+(1+\beta^2)\lambda+\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1+\beta)^2\lambda)}}{2\beta\lambda} > \frac{\kappa^2+(1+\beta^2)\lambda+\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1-\beta)^2\lambda)}}{2\beta\lambda} = \frac{\kappa^2+(1-\beta+\beta^2)\lambda}{\beta\lambda} > \frac{1-\beta+\beta^2}{\beta} = \frac{1}{\beta} + \beta - 1 > 1$. However, the given solution is inside the unit circle as $\frac{\kappa^2+(1+\beta^2)\lambda-\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1+\beta)^2\lambda)}}{2\beta\lambda} > \frac{\kappa^2+(1+\beta^2)\lambda-\sqrt{(\kappa^2+(1+\beta)^2\lambda)(\kappa^2+(1+\beta)^2\lambda)}}{2\beta\lambda} = -1$, $\frac{\kappa^2+(1+\beta^2)\lambda-\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1+\beta)^2\lambda)}}{2\beta\lambda} < \frac{\kappa^2+(1+\beta^2)\lambda-\sqrt{(\kappa^2+(1-\beta)^2\lambda)(\kappa^2+(1-\beta)^2\lambda)}}{2\beta\lambda} = 1$.

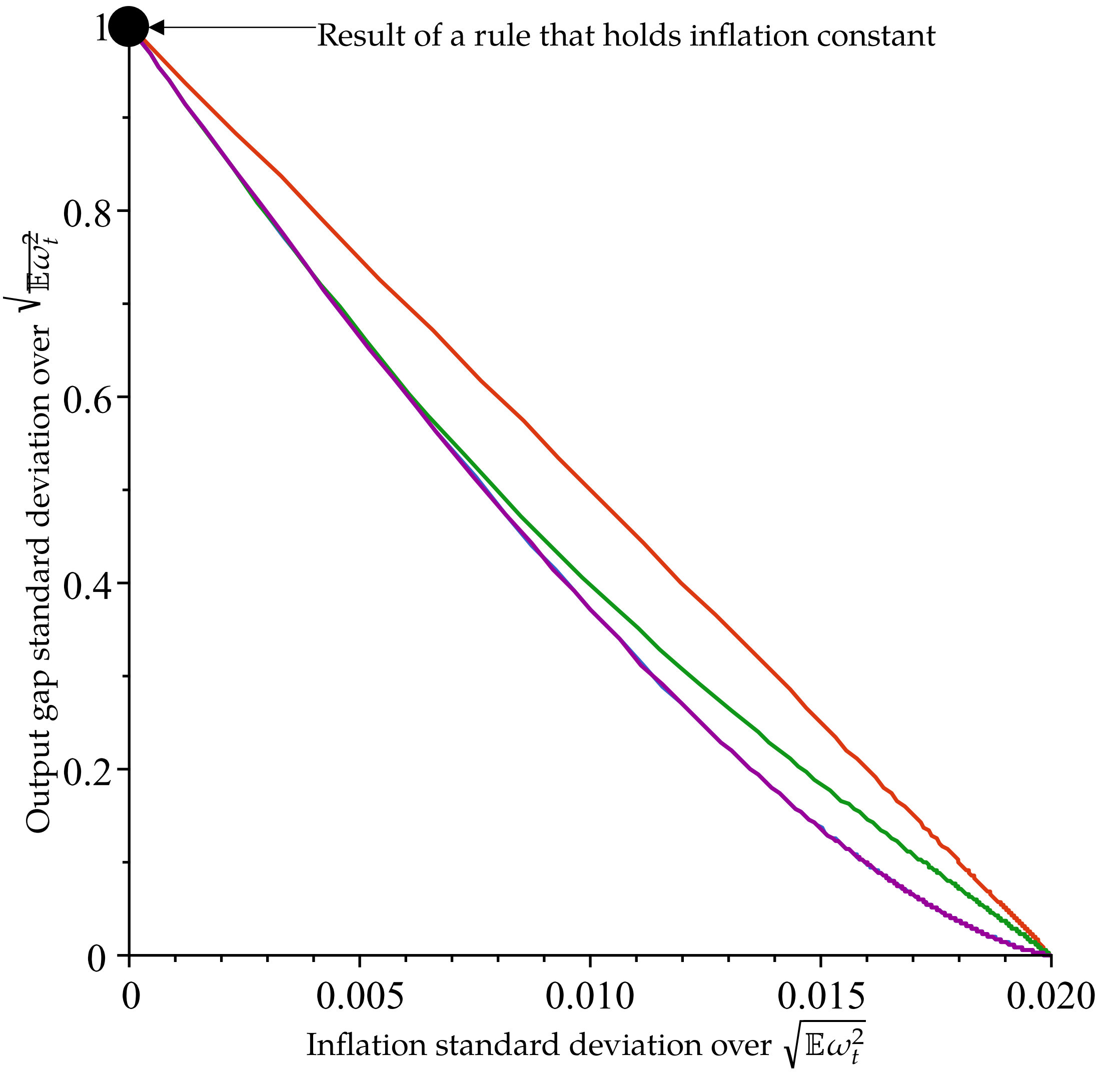

Figure 1: Policy frontiers (values attained by varying λ). $\beta = 0.99$, $\kappa = 0.02$.
Purple: Unconditionally optimal policy, equivalent to ARMA$(1, 1)$ policy.
Blue (hidden behind purple): Timeless optimal solution.
Red: Policy just responding to current shocks, equivalent to discretion.
Green: Policy that responds to current and once lagged shocks.

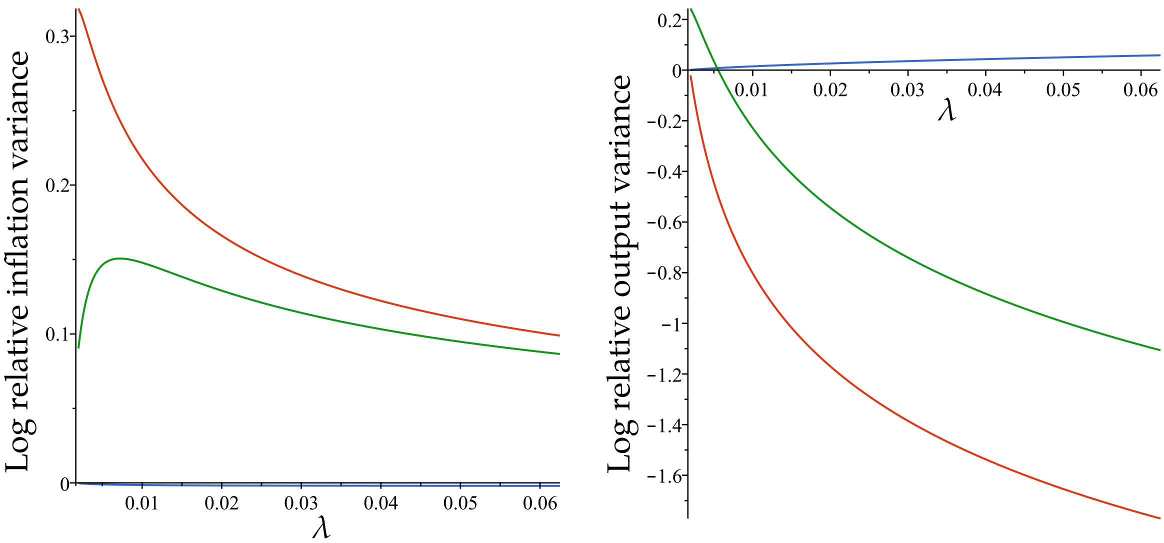

Figure 2: Logarithms of ratios of variance under a given policy to variance under unconditionally optimal policy. $\beta = 0.99$, $\kappa = 0.02$.
Blue: Timeless optimal solution.
Red: Policy just responding to current shocks, equivalent to discretion.
Green: Policy that responds to current and once lagged shocks.

To see the welfare attained by the other policies we have discussed, Figure 1 plots the policy frontiers attained by varying λ for each of the polices. In all cases, we follow Eggertsson & Woodford (2003) in setting $\beta = 0.99$ and $\kappa = 0.02$. The figure makes clear that the MA(1) policy (green) is a substantial improvement on the MA(0) (discretionary) policy (red). It also shows just how close Woodford's timeless perspective (1999)[34] (blue, hidden behind purple) comes to the unconditionally optimal policy. Figure 2 shows how these differences across policies are driven by λ, by plotting the logarithm of the ratio of variance under a given policy to the variance under unconditionally optimal policy. We allow λ to vary from 0.002 (the value obtained by a second order approximation to the consumer's utility with $\kappa = 0.02$, if the elasticity of substitution across goods equals 10) to $\frac{1}{16}$ (corresponding to an equal weight on annual inflation and the output gap). Both the MA(0) and the MA(1) policy generate too much inflation variance and too little variance in output, relative to the unconditionally optimal solution. However, if the central bank can feasibly respond to

[34] See Appendix A.8 for the derivation of this solution.

ω_t and ω_{t-1} they can probably also respond to π_{t-1}, which is enough to deliver the unconditional optimum.

Even in larger models, optimal inflation dynamics appear to be well approximated by an ARMA process with relatively few MA terms. Figure 3 shows the dynamics of observed and optimal inflation in the Justiniano, Primiceri & Tambalotti (2013) model. (This is a medium-scale New Keynesian DSGE model broadly similar to the model of Smets & Wouters (2007).) While actual inflation is highly persistent, with the same shocks hitting the economy, optimal inflation is far less persistent, with the sample autocorrelation essentially insignificant at 95% after four lags.

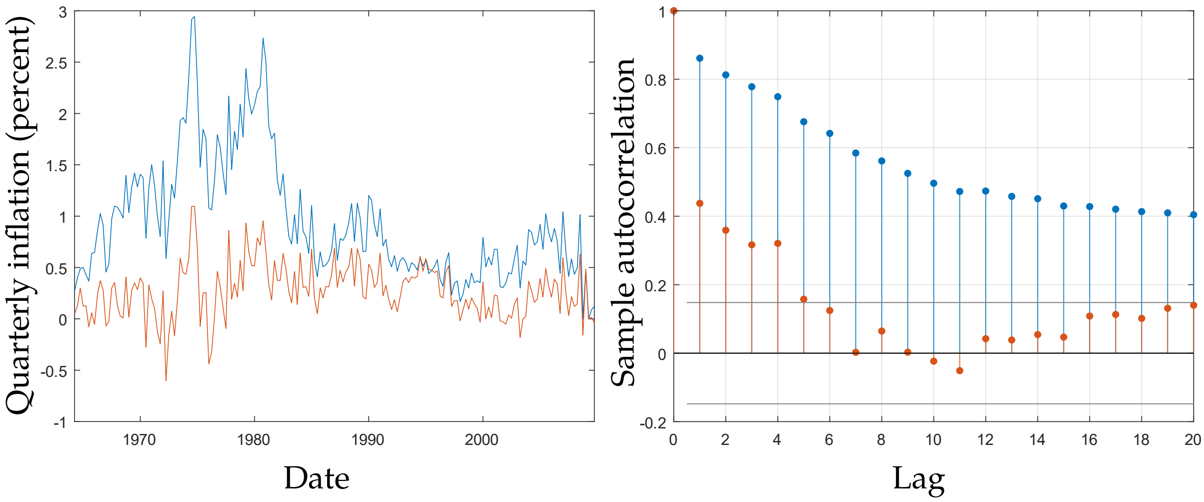

Figure 3: Behaviour of realised inflation (blue) and optimal inflation (red) in the Justiniano, Primiceri & Tambalotti (2013) model.
Left panel shows the timeseries. Right panel shows their sample autocorrelation.

Note that for any $\rho \in (-1,1)$, the solution for optimal inflation has a multiple shock, $\mathrm{ARMA}(1,\infty)$ representation of the form:

$$\pi_t - \pi = \rho(\pi_{t-1} - \pi) + \sum_{k=0}^{\infty} \sum_{n=1}^{N} \theta_{n,k}^{(\rho)} \varepsilon_{n,t-k},$$

where $\varepsilon_{1,t}, \dots, \varepsilon_{N,t}$ are the model's structural shocks. We can approximate this process by truncating the MA terms at some point, e.g. by considering the multiple shock $\mathrm{ARMA}(1,K)$ process:

$$\pi_t^{(K)} - \pi = \rho\left(\pi_{t-1}^{(K)} - \pi\right) + \sum_{k=0}^{K} \sum_{n=1}^{N} \theta_{n,k}^{(\rho)} \varepsilon_{n,t-k}.$$

In Figure 4 we plot the proportion of the variance of optimal inflation that is explained by this truncated process for $K = 0, \ldots, 16$, and $\rho \in \{0, 0.61\}$.[35] A multiple shock ARMA(1,1) process already explains over 90% of the variance of optimal inflation, while a multiple shock ARMA(1,2) explains over 95%. Thus, optimal inflation in plausible models can be well approximated by relatively simple inflation dynamics.

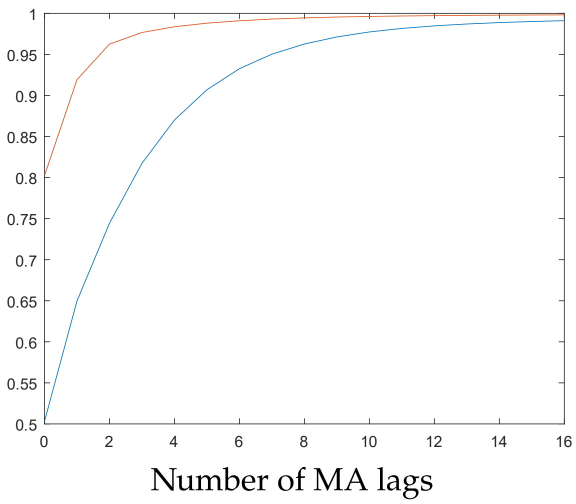

Figure 4: Proportion of the variance of optimal inflation in the Justiniano, Primiceri & Tambalotti (2013) model explained by truncating the number of MA lags. Blue: $\rho = 0$. Red: $\rho = 0.61$.

TODO (MORE, BIGGER MODELS, ETC)

5 Empirical support

In this section we examine whether it is possible that the US Federal Reserve is already using a simple rule of our form. Without the restriction to simple rules, the answer to this would be trivial. Since the rule of equation (6) is compatible with arbitrary inflation dynamics, any observed dynamics can be explained by such a rule.

[35] $\rho = 0.61$ is the value of ρ that minimises the variance of $\sum_{k=0}^{\infty} \sum_{n=1}^{N} \theta_{n,k}^{(\rho)} \varepsilon_{n,t-k}$. I.e. it is the value of ρ that would be estimated by OLS using an infinite sample of observations from optimal inflation.

A natural approach is to take inspiration from the inflation dynamics estimated in reduced form work. Dotsey, Fujita & Stark (2018) find that an IMA$(1,1)$ model for inflation forecasts well, and Stock & Watson (2009) note that this is well approximated by an ARMA$(1,1)$.

Thus, as a first experiment we look at the performance of rules which ensure inflation follows an ARMA$(1,1)$ (as stationarity is convenient),

TODO WRITE UP MODEL ESTIMATED ON MONTHLY DATA

TODO MORE! (Or maybe just cut this whole section completely)

6 The zero lower bound

TODO (NOTE: Helps avoid the ZLB, longer rates are less likely to hit zero.)

TODO Rules based on first differences of $i_t - r_t$

TODO Using perpetuities

7 Conclusion

TODO

8 References

Adão, Bernardino, Isabel Correia & Pedro Teles. 2011. 'Unique Monetary Equilibria with Interest Rate Rules'. *Review of Economic Dynamics* 14 (3) (July): 432–442.

Ang, Andrew, Geert Bekaert & Min Wei. 2007. 'Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?' *Journal of Monetary Economics* 54 (4) (May 1): 1163–1212.

Ascari, Guido & Tiziano Ropele. 2009. 'Trend Inflation, Taylor Principle, and Indeterminacy'. *Journal of Money, Credit and Banking* 41 (8): 1557–1584.

Atkeson, Andrew & Lee E Ohanian. 2001. 'Are Phillips Curves Useful for Forecasting Inflation?' Edited by Edward J Green & Richard M Todd. *Federal Reserve Bank of Minneapolis Quarterly Review* (Winter 2001): 12.

Auclert, Adrien. 2019. 'Monetary Policy and the Redistribution Channel'. *American Economic Review* 109 (6) (June): 2333–2367.

Bilbiie, Florin. 2008. 'Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic'. *Journal of Economic Theory* 140 (1) (May 1): 162–196.

———. 2019. *Monetary Policy and Heterogeneity: An Analytical Framework*. 2019 Meeting Papers. Society for Economic Dynamics.

Branch, William A. & Bruce McGough. 2010. 'Dynamic Predictor Selection in a New Keynesian Model with Heterogeneous Expectations'. *Journal of Economic Dynamics and Control* 34 (8) (August 1): 1492–1508.

———. 2018. 'Chapter 1 - Heterogeneous Expectations and Micro-Foundations in Macroeconomics'. In *Handbook of Computational Economics*, edited by Cars Hommes & Blake LeBaron, 4:3–62. Handbook of Computational Economics. Elsevier.

Christiano, Lawrence J. & Yuta Takahashi. 2018. 'Discouraging Deviant Behavior in Monetary Economics'. *NBER Working Paper Series* No. 24949.

Cochrane, John H. 2011. 'Determinacy and Identification with Taylor Rules'. *Journal of Political Economy* 119 (3): 565–615.

———. 2017. 'The Grumpy Economist: Target the Spread'. *The Grumpy Economist*.

———. 2018. 'The Zero Bound, Negative Rates, and Better Rules' (March 2): 27.

Cúrdia, Vasco, Andrea Ferrero, Ging Cee Ng & Andrea Tambalotti. 2015. 'Has U.S. Monetary Policy Tracked the Efficient Interest Rate?' *Journal of Monetary Economics* 70: 72–83.

Damjanovic, Tatiana, Vladislav Damjanovic & Charles Nolan. 2008. 'Unconditionally Optimal Monetary Policy'. *Journal of Monetary Economics* 55 (3) (April 1): 491–500.

Doepke, Matthias & Martin Schneider. 2006. 'Inflation and the Redistribution of Nominal Wealth'. *Journal of Political Economy* 114 (6) (December): 1069–1097.

Dotsey, Michael, Shigeru Fujita & Tom Stark. 2018. 'Do Phillips Curves Conditionally Help to Forecast Inflation?' *International Journal of Central Banking*: 50.

Dotsey, Michael & Andreas Hornstein. 2006. 'Implementation of Optimal Monetary Policy'. *Economic Quarterly*: 113–133.

Dowd, Kevin. 1994. 'A Proposal to End Inflation'. *The Economic Journal* 104 (425): 828–840.

Eggertsson, Gauti B. & Michael Woodford. 2003. 'The Zero Bound on Interest Rates and Optimal Monetary Policy'. *Brookings Papers on Economic Activity* 34 (1): 139–235.

Evans, George W. & Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Frontiers of Economic Research. Princeton and Oxford: Princeton University Press.

———. 2006. 'Monetary Policy, Expectations and Commitment'. *The Scandinavian Journal of Economics* 108 (1): 15–38.

Evans, George W. & Bruce McGough. 2010. 'Implementing Optimal Monetary Policy in New-Keynesian Models with Inertia'. *The B.E. Journal of Macroeconomics* 10 (1).

Farmer, Roger E.A., Vadim Khramov & Giovanni Nicolò. 2015. 'Solving and Estimating Indeterminate DSGE Models'. *Journal of Economic Dynamics and Control* 54 (May 1): 17–36.

Gali, Jordi, J. David Lopez-Salido & Javier Valles. 2004. 'Rule-of-Thumb Consumers and the Design of Interest Rate Rules'. *Journal of Money, Credit, and Banking* 36 (4) (July 28): 739–763.

Galí, Jordi, J. David López-Salido & Javier Vallés. 2007. 'Understanding the Effects of Government Spending on Consumption'. *Journal of the European Economic Association* 5 (1) (March 1): 227–270.

Gürkaynak, Refet S., Brian Sack & Jonathan H. Wright. 2010. 'The TIPS Yield Curve and Inflation Compensation'. *American Economic Journal: Macroeconomics* 2 (1): 70–92.

Hall, Robert E & Ricardo Reis. 2016. *Achieving Price Stability by Manipulating the Central Bank's Payment on Reserves*. Working Paper. National Bureau of Economic Research.

Hansen, Gary D. 1985. 'Indivisible Labor and the Business Cycle'. *Journal of Monetary Economics* 16 (3) (November 1): 309–327.

Hellwig, Martin F. 1980. 'On the Aggregation of Information in Competitive Markets'. *Journal of Economic Theory* 22 (3) (June 1): 477–498.

Hetzel, Robert L. 1990. 'Maintaining Price Stability: A Proposal'. *Economic Review of the Federal Reserve Bank of Richmond* (March): 3.

Holden, Tom D. 2019. *Existence and Uniqueness of Solutions to Dynamic Models with Occasionally Binding Constraints*. EconStor Preprints. ZBW - Leibniz Information Centre for Economics.

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

———. 2015. 'Monetary Policy, Bond Risk Premia, and the Economy'. *Journal of Monetary Economics* 76 (November 1): 124–140.

Justiniano, Alejandro, Giorgio E. Primiceri & Andrea Tambalotti. 2013. 'Is There a Trade-Off between Inflation and Output Stabilization?' *American Economic Journal: Macroeconomics* 5 (2) (April): 1–31.

Kaplan, Greg, Benjamin Moll & Giovanni L. Violante. 2018. 'Monetary Policy According to HANK'. *American Economic Review* 108 (3) (March): 697–743.

Khan, Hashmat. 2005. 'Price-Setting Behaviour, Competition, and Markup Shocks in the New Keynesian Model'. *Economics Letters* 87 (3) (June 1): 329–335.

Khan, Hashmat, Louis Phaneuf & Jean Gardy Victor. 2019. 'Rules-Based Monetary Policy and the Threat of Indeterminacy When Trend Inflation Is Low'. *Journal of Monetary Economics* (March): S0304393219300479.

Lou, Youcheng, Sahar Parsa, Debraj Ray, Duan Li & Shouyang Wang. 2019. 'Information Aggregation in a Financial Market with General Signal Structure'. *Journal of Economic Theory* 183 (September 1): 594–624.

Lubik, Thomas A., Christian Matthes & Elmar Mertens. 2019. *Indeterminacy and Imperfect Information*. Federal Reserve Bank of Richmond Working Papers.

Manea, Cristina. 2019. 'Collateral-Constrained Firms and Monetary Policy': 66.

McGough, Bruce, Glenn D. Rudebusch & John C. Williams. 2005. 'Using a Long-Term Interest Rate as the Monetary Policy Instrument'. *Journal of Monetary Economics* 52 (5) (July 1): 855–879.

McKay, Alisdair, Emi Nakamura & Jón Steinsson. 2017. 'The Discounted Euler Equation: A Note'. *Economica* 84 (336): 820–831.

McLeay, Michael & Silvana Tenreyro. 2019. 'Optimal Inflation and the Identification of the Phillips Curve'. In . NBER Chapters. National Bureau of Economic Research, Inc.

Natvik, Gisle James. 2009. 'Government Spending and the Taylor Principle'. *Journal of Money, Credit and Banking* 41 (1): 57–77.

Smets, Frank & Rafael Wouters. 2007. 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach'. *American Economic Review* 97 (3) (June): 586–606.

Stock, James & Mark W. Watson. 2009. 'Phillips Curve Inflation Forecasts'. In *Understanding Inflation and the Implications for Monetary Policy*, edited by Jeffrey

Fuhrer, Yolanda Kodrzycki, Jane Little & Giovanni Olivei, 99–202. Cambridge: MIT Press.

Sveen, Tommy & Lutz Weinke. 2005. 'New Perspectives on Capital, Sticky Prices, and the Taylor Principle'. *Journal of Economic Theory* 123 (1). Monetary Policy and Capital Accumulation (July 1): 21–39.

Svensson, Lars E. O & Michael Woodford. 2003. 'Implementing Optimal Policy through Inflation-Forecast Targeting' (June).

Woodford, Michael. 1999. 'Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?' In *New Challenges for Monetary Policy*. Jackson Hole, Wyoming: Federal Reserve Bank of Kansas City.

———. 2001. 'The Taylor Rule and Optimal Monetary Policy'. *The American Economic Review* 91 (2): 232–237.

———. 2003. *Interest and Prices. Foundations of a Theory of Monetary Policy*. Princeton University Press.

Online Appendix to: "Robust Real Rate Rules"

Tom D. Holden, Deutsche Bundesbank[36] 28/02/2022

Appendix A Proofs and supplemental results

A.1 Phillips curve based forecasting with ARMA(1,1) policy shocks

As before, we have the monetary rule:

$$i_t = r_t + \phi\pi_t + \zeta_t,$$

which combined with the Fisher equation gives:

$$\mathbb{E}_t\pi_{t+1} = \phi\pi_t + \zeta_t.$$

Suppose ζ_t follows the ARMA(1,1) process:

$$\zeta_t = \rho_\zeta\zeta_{t-1} + \varepsilon_{\zeta,t} + \theta_\zeta\varepsilon_{\zeta,t-1}, \qquad \varepsilon_{\zeta,t} \sim N(0,\sigma_\zeta^2)$$

with $\rho_\zeta, \theta_\zeta \in (-1,1)$. Then from matching coefficients, with $\phi > 1$ we have the unique solution:

$$\pi_t = -\frac{1}{\phi - \rho_\zeta}\left[\zeta_t + \frac{\theta_\zeta}{\phi}\varepsilon_{\zeta,t}\right].$$

Thus:

$$\pi_t - \rho_\zeta\pi_{t-1} = -\frac{1}{\phi - \rho_\zeta}\left(1 + \frac{\theta_\zeta}{\phi}\right)\left[\varepsilon_{\zeta,t} + \frac{\phi - \rho_\zeta}{\phi + \theta_\zeta}\theta_\zeta\varepsilon_{\zeta,t-1}\right],$$

so π_t also follows an ARMA(1,1) process. Suppose for now that $-\rho_\zeta \le \theta_\zeta$, which is likely to be satisfied in reality as we expect ρ_ζ to be large and positive, while θ_ζ should be close to zero. (For example, Dotsey, Fujita & Stark (2018) find that an IMA(1,1) model fits inflation well, in which case $-\rho_\zeta = -1 < \theta_\zeta$ as required.) Then $0 < \frac{\phi-\rho_\zeta}{\phi+\theta_\zeta} < 1$, so $\left|\frac{\phi-\rho_\zeta}{\phi+\theta_\zeta}\theta_\zeta\right| < 1$ meaning the process for inflation is invertible. With inflation

[36] Wilhelm-Epstein-Straße 14, 60431, Frankfurt am Main, Germany. E-mail: thomas.holden@gmail.com.

following an invertible linear process, the full-information optimal forecast of π_{t+1} is a linear combination of $\pi_t, \pi_{t-1}, \ldots$. In particular, as before x_t is not useful.

In the unlikely case in which $-\rho_\zeta > \theta_\zeta$, of if the forecaster's information set $\mathcal{I}_t$ is smaller than $\{\pi_t, x_t, \pi_{t-1}, x_{t-1}, \ldots\}$,[37] then x_t may contain some useful information. Combining the solution for inflation with the Phillips curve:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + \kappa \omega_t,$$

gives:

$$\begin{aligned} x_t &= -\frac{1}{\kappa}\left[\frac{1-\beta\rho_\zeta}{\phi-\rho_\zeta}\left(\zeta_t + \frac{\theta_\zeta}{\phi}\varepsilon_{\zeta,t}\right) - \beta\frac{\theta_\zeta}{\phi}\varepsilon_{\zeta,t}\right] - \omega_t \\ &= \frac{1}{\kappa}\left[(1-\beta\rho_\zeta)\pi_t + \beta\frac{\theta_\zeta}{\phi}\varepsilon_{\zeta,t}\right] - \omega_t. \end{aligned}$$

In this case, it is possible that $\mathbb{E}[\pi_{t+1}|\mathcal{I}_t] \neq \mathbb{E}[\pi_{t+1}|\mathcal{I}_{t-1}, \pi_t]$ as x_t provides an independent signal about $\varepsilon_{\zeta,t}$.

There are two important special cases. If $\omega_t = 0$, and the forecaster knows this, then:

$$\varepsilon_{\zeta,t} = \frac{\phi}{\beta\theta_\zeta}\left[\kappa x_t - (1-\beta\rho_\zeta)\pi_t\right],$$

so:

$$\zeta_t = -\left(\phi - \frac{1}{\beta}\right)\pi_t - \frac{\kappa}{\beta}x_t,$$

which enables the forecaster to form the full-information optimal forecast:

$$\mathbb{E}_t\pi_{t+1} = -\frac{1}{\phi-\rho_\zeta}\left(\rho_\zeta\zeta_t + \theta_\zeta\varepsilon_{\zeta,t}\right) = \frac{1}{\beta}\left(\pi_t - \kappa x_t\right).$$

(This formula also follows immediately from the Phillips curve.) Note that the output gap has what Dotsey, Fujita & Stark (2018) call the "wrong" sign, meaning Phillips curve based forecasting regressions may have surprising results. However, in the general case in which ω_t has positive variance, then output's signal about $\varepsilon_{\zeta,t}$ will be polluted by the noise from ω_t, making it much less informative. Indeed, with ϕ large, as we expect, then $\frac{\theta_\zeta}{\phi}\varepsilon_{\zeta,t}$ will have low variance, making it more likely that it is drowned out by the noise from ω_t.

[37] We nonetheless assume that π_t and x_t are in $\mathcal{I}_t$.

The second important special case is when $\varepsilon_{\zeta,t} = 0$, and again the forecaster knows this. In this case, much as in the main text:

$$\mathbb{E}_t \pi_{t+1} = \rho_\zeta \pi_t - \frac{1}{\phi - \rho_\zeta}\left(1 + \frac{\theta_\zeta}{\phi}\right)\left[\mathbb{E}_t \varepsilon_{\zeta,t+1} + \frac{\phi - \rho_\zeta}{\phi + \theta_\zeta}\theta_\zeta \varepsilon_{\zeta,t}\right] = \rho_\zeta \pi_t,$$

so x_t is unhelpful.

The general case will inherit aspects of these two special cases, as well as the case in which π_t's stochastic process was invertible. Inflation and its lags will certainly help forecast inflation, but the output gap may also provide a little extra information, possibly with the "wrong" sign.

A.2 Robustness to non-unit responses to real interest rates

Suppose that the central bank is unable to respond with a precise unit coefficient to real interest rates, so instead follows the monetary rule:

$$i_t = (1 + \gamma) r_t + \phi \pi_t + \zeta_t,$$

where $\gamma \in \mathbb{R}$ is some small value giving the departure from unit responses.

For simplicity, suppose the rest of the model takes the same form as in Section 1, with:

$$x_t = \delta \mathbb{E}_t x_{t+1} - \varsigma (r_t - n_t),$$
$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + \kappa \omega_t,$$
$$i_t = r_t + \mathbb{E}_t \pi_{t+1}.$$

We suppose $\phi > 1$, but do not make any assumptions on the signs of $\delta, \beta, \kappa, \varsigma, \gamma$, beyond assuming that $\varsigma \neq 0$ (so monetary policy has some effect on the output gap) and $\kappa \neq 0$ (so monetary policy has some effect on inflation, via the output gap).

Combining the monetary rule with the Fisher equation gives:

$$\mathbb{E}_t \pi_{t+1} = \gamma r_t + \phi \pi_t + \zeta_t,$$

so:

$$r_t = \frac{1}{\gamma}(\mathbb{E}_t \pi_{t+1} - \phi \pi_t - \zeta_t),$$

meaning:

$$x_t = \delta \mathbb{E}_t x_{t+1} - \frac{\varsigma}{\gamma}(\mathbb{E}_t \pi_{t+1} - \phi \pi_t) + \varsigma n_t + \frac{\varsigma}{\gamma}\zeta_t.$$

Then, since:

$$\mathbb{E}_t \pi_{t+1} = \frac{1}{\beta}\pi_t - \frac{\kappa}{\beta}x_t - \frac{\kappa}{\beta}\omega_t,$$

we have that:

$$\mathbb{E}_t x_{t+1} = \left(\frac{1}{\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}\right)x_t - \frac{\varsigma}{\delta\gamma}\left(\phi - \frac{1}{\beta}\right)\pi_t - \frac{\varsigma}{\delta\gamma}\left(\gamma n_t + \zeta_t + \frac{\kappa}{\beta}\omega_t\right).$$

Woodford (2003) (Addendum to Chapter 4, Proposition C.1) proves that this model is determinate if and only if both eigenvalues of the matrix:

$$M := \begin{bmatrix} \frac{1}{\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta} & -\frac{\varsigma}{\delta\gamma}\left(\phi - \frac{1}{\beta}\right) \\ -\frac{\kappa}{\beta} & \frac{1}{\beta} \end{bmatrix}$$

are outside of the unit circle, which in turn is proven to hold if and only if EITHER: Case I: $1 < \det M$, $0 < 1 + \det M - \operatorname{tr} M$, and $0 < 1 + \det M + \operatorname{tr} M$, OR Case II: $0 > 1 + \det M - \operatorname{tr} M$, and $0 > 1 + \det M + \operatorname{tr} M$. Note:

$$\det M = \frac{1}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}\phi,$$
$$\operatorname{tr} M = \frac{1}{\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta} + \frac{1}{\beta}.$$

Thus, Case I requires:

$$1 < \det M = \frac{1}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}\phi,$$
$$0 < 1 + \det M - \operatorname{tr} M = \frac{(1-\beta)(1-\delta)}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}(\phi - 1),$$
$$\text{and } 0 < 1 + \det M + \operatorname{tr} M = \frac{(1+\beta)(1+\delta)}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}(1 + \phi).$$

And Case II requires:

$$0 > 1 + \det M - \operatorname{tr} M = \frac{(1-\beta)(1-\delta)}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}(\phi - 1),$$
$$\text{and } 0 > 1 + \det M + \operatorname{tr} M = \frac{(1+\beta)(1+\delta)}{\beta\delta} - \frac{\varsigma\kappa}{\gamma\beta\delta}(1 + \phi).$$

To see when these conditions are satisfied, first suppose that $\frac{\varsigma\kappa}{\gamma\beta\delta} < 0$, so $\frac{\varsigma\kappa}{\gamma\beta\delta} = -\frac{|\varsigma\kappa|}{|\gamma||\beta\delta|}$. Then if γ is sufficiently small in magnitude, it is immediately clear that all three conditions of Case I are satisfied, since $\phi > 0$, $\phi - 1 > 0$ and $1 + \phi > 0$. In particular, in this case we need:

$$|\gamma| < |\varsigma\kappa| \min \left\{ \begin{array}{c} \dfrac{\phi}{\max\{0, -(\operatorname{sign}(\beta\delta) - |\beta\delta|)\}}, \\ \dfrac{\phi - 1}{\max\{0, -(\operatorname{sign}(\beta\delta))(1-\beta)(1-\delta)\}}, \\ \dfrac{1+\phi}{\max\{0, -(\operatorname{sign}(\beta\delta))(1+\beta)(1+\delta)\}} \end{array} \right\}.$$

Alternatively, suppose that $\frac{\varsigma\kappa}{\gamma\beta\delta} > 0$, so $\frac{\varsigma\kappa}{\gamma\beta\delta} = \frac{|\varsigma\kappa|}{|\gamma||\beta\delta|}$. Then, similarly, if γ is sufficiently small in magnitude, both conditions of Case II are satisfied, since $\phi - 1 > 0$ and $1 + \phi > 0$. In particular, in this case we need:

$$|\gamma| < |\varsigma\kappa| \min \left\{ \begin{array}{c} \dfrac{\phi - 1}{\max\{0, (\operatorname{sign}(\beta\delta))(1-\beta)(1-\delta)\}}, \\ \dfrac{1+\phi}{\max\{0, (\operatorname{sign}(\beta\delta))(1+\beta)(1+\delta)\}} \end{array} \right\}.$$

Thus, it is always sufficient for determinacy that:

$$|\gamma| < |\varsigma\kappa| \min \left\{ \begin{array}{c} \dfrac{\phi}{\max\{0, -(\operatorname{sign}(\beta\delta) - |\beta\delta|)\}}, \\ \dfrac{\phi - 1}{|(1-\beta)(1-\delta)|}, \\ \dfrac{1+\phi}{|(1+\beta)(1+\delta)|} \end{array} \right\}.$$

Since the right-hand side is strictly positive, there is a positive measure of γ for which we have determinacy.

A.3 Real-time learning of Phillips curve coefficients

We start by assuming that the central bank knows the Phillips curve coefficients. A close examination of this case will lead to a natural learning scheme for when the central bank does not know these coefficients.

As in the main text, suppose the central bank is using the rule:

$$i_t = r_t + \phi_\pi \pi_t + \phi_x \Big[x_t - \kappa^{-1} \big[\pi_t - \tilde{\beta}(1-\varrho_\pi)\mathbb{E}_t \pi_{t+1} - \tilde{\beta}\varrho_\pi \pi_{t-1} \big] \Big] + \zeta_t,$$

and that the model also contains the Phillips curve:

$$\pi_t = \tilde{\beta}(1-\varrho_\pi)\mathbb{E}_t \pi_{t+1} + \tilde{\beta}\varrho_\pi \pi_{t-1} + \kappa x_t + \kappa \omega_t,$$

and the Fisher equation:

$$i_t = r_t + \mathbb{E}_t \pi_{t+1}.$$

We suppose that ζ_t follows the ARMA(1,1) process:

$$\zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_{\zeta,t} + \theta_\zeta \varepsilon_{\zeta,t-1}, \qquad \varepsilon_{\zeta,t} \sim N(0,\sigma_\zeta^2),$$

with $\rho_\zeta, \theta_\zeta \in (-1,1)$, and for simplicity, we suppose that $\omega_t = \varepsilon_{\omega,t}$, where $\varepsilon_{\omega,t} \sim N(0,\sigma_\omega^2)$.

From combining all the above equations, we have that if $\phi_\pi > 1$, there is a unique solution with:

$$\pi_t = -\frac{1}{\phi_\pi - \rho_\zeta}\left[\zeta_t + \frac{\theta_\zeta}{\phi_\pi}\varepsilon_{\zeta,t}\right] + \frac{\phi_x}{\phi_\pi}\varepsilon_{\omega,t}.$$

Thus, if we define:

$$\begin{aligned}
m_0 &:= \frac{\sigma_\zeta^2}{\kappa(\phi_\pi - \rho_\zeta)}\left[\tilde{\beta}(1-\varrho_\pi)(\rho_\zeta + \theta_\zeta) - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\right],\\
m_1 &:= \frac{\sigma_\zeta^2}{\kappa(\phi_\pi - \rho_\zeta)}\left[\left[\tilde{\beta}(1-\varrho_\pi)\rho_\zeta - 1\right](\rho_\zeta + \theta_\zeta) + \tilde{\beta}\varrho_\pi\left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\right],\\
m_2 &:= \frac{\sigma_\zeta^2}{\kappa(\phi_\pi - \rho_\zeta)}\left[\left[\tilde{\beta}(1-\varrho_\pi)\rho_\zeta - 1\right]\rho_\zeta + \tilde{\beta}\varrho_\pi\right](\rho_\zeta + \theta_\zeta),
\end{aligned}$$

then by the Phillips curve $m_0 = \mathbb{E}x_t\varepsilon_{\zeta,t}$, $m_1 = \mathbb{E}x_t\varepsilon_{\zeta,t-1}$ and $m_2 = \mathbb{E}x_t\varepsilon_{\zeta,t-2}$. Also note that:

$$\begin{aligned}
\kappa &= \frac{\sigma_\zeta^2}{\phi_\pi - \rho_\zeta}\frac{\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2}{\rho_\zeta\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)m_0 - \left((\rho_\zeta+\theta_\zeta)m_1 - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)m_2\right)},\\
\tilde{\beta} &= \frac{\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)\left(m_0 - (\rho_\zeta m_1 - m_2)\right) - \frac{\phi_\pi + \theta_\zeta}{(\rho_\zeta+\theta_\zeta)\phi_\pi}\left((\rho_\zeta+\theta_\zeta)m_1 - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)m_2\right)}{\rho_\zeta\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)m_0 - \left((\rho_\zeta+\theta_\zeta)m_1 - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)m_2\right)},\\
\varrho_\pi &= -\frac{\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)(\rho_\zeta m_1 - m_2)}{\left(\rho_\zeta + \theta_\zeta - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)\left(m_0 - (\rho_\zeta m_1 - m_2)\right) - \frac{\phi_\pi + \theta_\zeta}{(\rho_\zeta+\theta_\zeta)\phi_\pi}\left((\rho_\zeta+\theta_\zeta)m_1 - \left(1+\frac{\theta_\zeta}{\phi_\pi}\right)m_2\right)}.
\end{aligned}$$

In other words, once the central bank knows m_0, m_1 and m_2 they can infer the parameters of the Phillips curve from the known properties of their monetary rule and monetary shock. This is essentially an instrumental variables regression. We are using $\varepsilon_{\zeta,t}$, $\varepsilon_{\zeta,t-1}$ and $\varepsilon_{\zeta,t-2}$ as instruments for $\mathbb{E}_t\pi_{t+1}$, π_t and π_{t-1} in a regression of the output gap on those variables. This works as long as $\theta_\zeta \neq 0$, else $\mathbb{E}_t\pi_{t+1}$ and π_t are colinear.

If the central bank does not know the true values of κ, $\tilde{\beta}$ and ϱ_π, we suppose they dynamically update estimates of m_0, m_1 and m_2 using the following decreasing gain learning rules (for $t > 0$):

$$m_{0,t} = m_{0,t-1} + t^{-1}\big(x_t \varepsilon_{\zeta,t} - m_{0,t-1}\big),$$
$$m_{1,t} = m_{1,t-1} + t^{-1}\big(x_t \varepsilon_{\zeta,t-1} - m_{1,t-1}\big),$$
$$m_{2,t} = m_{2,t-1} + t^{-1}\big(x_t \varepsilon_{\zeta,t-2} - m_{2,t-1}\big),$$

where $\iota \in (0,1]$ is a gain parameter. Then they can use the monetary rule:

$$i_t = r_t + \phi_\pi \pi_t + \phi_x [x_t + q_{1,t-1}\mathbb{E}_t \pi_{t+1} + q_{0,t-1}\pi_t + q_{-1,t-1}\pi_{t-1}] + \zeta_t,$$

where:

$$q_{1,t} := \frac{\phi_\pi - \rho_\zeta}{\sigma_\zeta^2} \frac{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right) m_{0,t} - \frac{\phi_\pi + \theta_\zeta}{(\rho_\zeta + \theta_\zeta)\phi_\pi}\left((\rho_\zeta + \theta_\zeta) m_{1,t} - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right) m_{2,t}\right)}{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2},$$

$$q_{0,t} := -\frac{\phi_\pi - \rho_\zeta}{\sigma_\zeta^2} \frac{\rho_\zeta\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right) m_{0,t} - \left((\rho_\zeta + \theta_\zeta) m_{1,t} - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right) m_{2,t}\right)}{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2},$$

$$q_{-1,t} := -\frac{\phi_\pi - \rho_\zeta}{\sigma_\zeta^2} \frac{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)(\rho_\zeta m_{1,t} - m_{2,t})}{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2}.$$

This is reasonable, as if $m_{0,t-1} \approx m_0$, $m_{1,t-1} \approx m_1$ and $m_{2,t-1} \approx m_2$ then $q_{1,t-1} \approx \kappa^{-1}\tilde{\beta}(1 - \varrho_\pi)$, $q_{0,t-1} \approx -\kappa^{-1}$ and $q_{-1,t-1} \approx \kappa^{-1}\tilde{\beta}\varrho_\pi$, so this monetary rule is approximately the same as the full information one previously considered. Using lagged estimates ($q_{1,t-1}$ not $q_{1,t}$ etc.) in the monetary rule reflects central bank information (processing) delays and simplifies the model's solution. It is also a common assumption in the reduced form learning literature (Evans & Honkapohja 2001).

With the new monetary rule, the model is no-longer linear. As a result, the exact solution is analytically intractable. However, we are only really interested in asymptotic dynamics. If $m_{0,t} \to m_0$, $m_{1,t} \to m_1$ and $m_{2,t} \to m_2$ as $t \to \infty$ then we know the asymptotic solution will be the stable full information one we found previously. We will analyse the system's behaviour with help from the stochastic approximation tools frequently used in the reduced form learning literature (Evans & Honkapohja

2001). These tools only require a zeroth order approximation in t^{-1} to the dynamics of x_t and π_t.[38] Intuitively, this is because x_t (hence π_t) enters the law of motion for $m_{0,t}$, $m_{1,t}$ and $m_{2,t}$ multiplied by t^{-1}, so a zeroth order approximation to the dynamics of x_t and π_t in t^{-1} delivers a first order approximation to the dynamics of $m_{0,t}$, $m_{1,t}$ and $m_{2,t}$ in t^{-1}.

We conjecture a time-varying coefficients solution with:

$$\pi_t = A_{t-1}\zeta_t + B_{t-1}\varepsilon_{\zeta,t} + C_{t-1}\varepsilon_{\omega,t} + D_{t-1}\pi_{t-1} + O(t^{-1}),$$

where we conjecture $A_t = A_{t-1} + O(t^{-1})$, $B_t = B_{t-1} + O(t^{-1})$, $C_t = C_{t-1} + O(t^{-1})$ and $D_t = D_{t-1} + O(t^{-1})$. Substituting this into the monetary rule, Fisher equation and Phillips curve implies:

$$\begin{aligned}&\left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t-1}\right]A_t\left(\rho_\zeta\zeta_t + \theta_\zeta\varepsilon_{\zeta,t}\right)\\ &\qquad = \left[\phi_\pi + \phi_x\kappa^{-1} + \phi_x q_{0,t-1} - \left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t-1}\right]D_t\right]\left[A_{t-1}\zeta_t\right.\\ &\qquad \left. + B_{t-1}\varepsilon_{\zeta,t} + C_{t-1}\varepsilon_{\omega,t} + D_{t-1}\pi_{t-1}\right] + \phi_x\left[q_{-1,t-1} - \kappa^{-1}\tilde{\beta}\varrho_\pi\right]\pi_{t-1} - \phi_x\varepsilon_{\omega,t}\\ &\qquad + \zeta_t + O(t^{-1}).\end{aligned}$$

Matching terms and using $A_t = A_{t-1} + O(t^{-1})$ and $D_t = D_{t-1} + O(t^{-1})$ then gives that:

$$\begin{aligned}&\left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\right]A_t\rho_\zeta\\ &\qquad = \left[\phi_\pi + \phi_x\kappa^{-1} + \phi_x q_{0,t} - \left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\right]D_t\right]A_t + 1\\ &\qquad + O(t^{-1}),\\ &\left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\right]A_t\theta_\zeta\\ &\qquad = \left[\phi_\pi + \phi_x\kappa^{-1} + \phi_x q_{0,t} - \left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\right]D_t\right]B_t + O(t^{-1}),\\ &0 = \left[\phi_\pi + \phi_x\kappa^{-1} + \phi_x q_{0,t} - \left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t-1}\right]D_t\right]C_t - \phi_x + O(t^{-1}),\\ &0 = \left[\phi_\pi + \phi_x\kappa^{-1} + \phi_x q_{0,t} - \left[1 + \phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\right]D_t\right]D_t + \phi_x\left[q_{-1,t} - \kappa^{-1}\tilde{\beta}\varrho_\pi\right]\\ &\qquad + O(t^{-1}).\end{aligned}$$

[38] Given certain regularity conditions on the higher order terms. These conditions will be satisfied here, at least providing we restrict $m_{0,t}$, $m_{1,t}$ and $m_{2,t}$ to a small enough open set around m_0, m_1 and m_2, using a so called projection facility.

The final equation has two roots, but we know we need to pick the one that gives $D_t \to 0$ as $\phi_x \to 0$. Now if $q_{0,t}$ is sufficiently close to q_0, then $\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t} > 0$, so:

$$D_t = \frac{(\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t}) - \sqrt{\begin{array}{c}(\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t})^2 \cdots \\ +4\phi_x\big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big]\big[q_{-1,t} - \kappa^{-1}\tilde{\beta}\varrho_\pi\big]\end{array}}}{2\big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big]} + O(t^{-1}),$$

and:

$$\begin{aligned}
A_t &= \Big[\big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big](D_t + \rho_\zeta) - (\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t})\Big]^{-1} + O(t^{-1}), \\
B_t &= \frac{\theta_\zeta\big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big]A_t}{\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t} - \big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big]D_t} + O(t^{-1}), \\
C_t &= \frac{\phi_x}{\phi_\pi + \phi_x \kappa^{-1} + \phi_x q_{0,t} - \big[1 + \phi_x \kappa^{-1}\tilde{\beta}(1-\varrho_\pi) - \phi_x q_{1,t}\big]D_t} + O(t^{-1}).
\end{aligned}$$

Since $q_{1,t} = q_{1,t-1} + O(t^{-1})$, $q_{0,t} = q_{0,t-1} + O(t^{-1})$ and $q_{-1,t} = q_{-1,t-1} + O(t^{-1})$, as required we have that $A_t = A_{t-1} + O(t^{-1})$, $B_t = B_{t-1} + O(t^{-1})$, $C_t = C_{t-1} + O(t^{-1})$ and $D_t = D_{t-1} + O(t^{-1})$.

Using this result again, we then have that:

$$\begin{aligned}
x_t = \kappa^{-1}\Big[&\big[1 - \tilde{\beta}(1-\varrho_\pi)(D_{t-1} + \rho_\zeta)\big]A_{t-1}\zeta_t \\
&+ \big[B_{t-1} - \tilde{\beta}(1-\varrho_\pi)\big(A_{t-1}\theta_\zeta + B_{t-1}D_{t-1}\big)\big]\varepsilon_{\zeta,t} \\
&+ \Big[\big[1 - \tilde{\beta}(1-\varrho_\pi)D_{t-1}\big]C_{t-1} - \kappa\Big]\varepsilon_{\omega,t} \\
&+ \Big[\big[1 - \tilde{\beta}(1-\varrho_\pi)D_{t-1}\big]D_{t-1} - \tilde{\beta}\varrho_\pi\Big]\pi_{t-1}\Big] + O(t^{-1}).
\end{aligned}$$

Plugging this into the law of motion for $m_{0,t}$, $m_{1,t}$ and $m_{2,t}$ gives a purely backward looking non-linear system in the endogenous states $m_{0,t}$, $m_{1,t}$, $m_{2,t}$ and π_t. This system is of the correct form to be analysed by the stochastic approximation results given in Evans & Honkapohja (2001).

To apply these results, first suppose that for all t, $m_{0,t} = \widehat{m}_0$, $m_{1,t} = \widehat{m}_1$ and $m_{2,t} = \widehat{m}_2$, for some values $\widehat{m}_0$, $\widehat{m}_1$ and $\widehat{m}_2$. Then $q_{1,t} = \hat{q}_1$, $q_{0,t} = \hat{q}_0$ and $q_{-1,t} = \hat{q}_{-1}$ for all t, where:

$$\hat{q}_1 := \frac{\phi_\pi - \rho_\zeta}{\sigma_\zeta^2} \frac{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)\widehat{m}_0 - \frac{\phi_\pi + \theta_\zeta}{(\rho_\zeta + \theta_\zeta)\phi_\pi}\left((\rho_\zeta + \theta_\zeta)\widehat{m}_1 - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\widehat{m}_2\right)}{\left(\rho_\zeta + \theta_\zeta - \left(1 + \frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2},$$

$$\hat{q}_0 := -\frac{\phi_\pi-\rho_\zeta}{\sigma_\zeta^2}\frac{\rho_\zeta\left(\rho_\zeta+\theta_\zeta-\left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)\widehat{m}_0-\left((\rho_\zeta+\theta_\zeta)\widehat{m}_1-\left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\widehat{m}_2\right)}{\left(\rho_\zeta+\theta_\zeta-\left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2},$$

$$\hat{q}_{-1} := -\frac{\phi_\pi-\rho_\zeta}{\sigma_\zeta^2}\frac{\left(\rho_\zeta+\theta_\zeta-\left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)(\rho_\zeta\widehat{m}_1-\widehat{m}_2)}{\left(\rho_\zeta+\theta_\zeta-\left(1+\frac{\theta_\zeta}{\phi_\pi}\right)\rho_\zeta\right)^2}.$$

Thus, for all t, $A_t = \hat{A}$, $B_t = \hat{B}$, $C_t = \hat{C}$ and $D_t = \widehat{D}$, where:

$$\widehat{D} = \frac{(\phi_\pi+\phi_x\kappa^{-1}+\phi_x\hat{q}_0) - \sqrt{\begin{array}{c}(\phi_\pi+\phi_x\kappa^{-1}+\phi_x\hat{q}_0)^2\cdots\\ +4\phi_x\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]\left[\hat{q}_{-1}-\kappa^{-1}\tilde{\beta}\varrho_\pi\right]\end{array}}}{2\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]},$$

and:

$$\hat{A} = \left[\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]\left(\widehat{D}+\rho_\zeta\right)-(\phi_\pi+\phi_x\kappa^{-1}+\phi_x\hat{q}_0)\right]^{-1},$$

$$\hat{B} = \frac{\theta_\zeta\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]\hat{A}}{\phi_\pi+\phi_x\kappa^{-1}+\phi_x\hat{q}_0-\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]\widehat{D}},$$

$$\hat{C} = \frac{\phi_x}{\phi_\pi+\phi_x\kappa^{-1}+\phi_x\hat{q}_0-\left[1+\phi_x\kappa^{-1}\tilde{\beta}(1-\varrho_\pi)-\phi_x\hat{q}_1\right]\widehat{D}}.$$

So:

$$\pi_t = \hat{A}\zeta_t + \hat{B}\varepsilon_{\zeta,t} + \hat{C}\varepsilon_{\omega,t} + \widehat{D}\pi_{t-1},$$

and:

$$\begin{aligned}
x_t &= \kappa^{-1}\Big[\left[1-\tilde{\beta}(1-\varrho_\pi)\left(\widehat{D}+\rho_\zeta\right)\right]\hat{A}\zeta_t + \left[\hat{B}-\tilde{\beta}(1-\varrho_\pi)\left(\hat{A}\theta_\zeta+\hat{B}\widehat{D}\right)\right]\varepsilon_{\zeta,t}\\
&\qquad + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\hat{C}-\kappa\right]\varepsilon_{\omega,t} + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\widehat{D}-\tilde{\beta}\varrho_\pi\right]\pi_{t-1}\Big]\\
&= \kappa^{-1}\Big[\left[1-\tilde{\beta}(1-\varrho_\pi)\left(\widehat{D}+\rho_\zeta\right)\right]\hat{A}\left[\rho_\zeta\left[\rho_\zeta\zeta_{t-2}+\varepsilon_{\zeta,t-1}+\theta_\zeta\varepsilon_{\zeta,t-2}\right]+\varepsilon_{\zeta,t}+\theta_\zeta\varepsilon_{\zeta,t-1}\right]\\
&\qquad + \left[\hat{B}-\tilde{\beta}(1-\varrho_\pi)\left(\hat{A}\theta_\zeta+\hat{B}\widehat{D}\right)\right]\varepsilon_{\zeta,t} + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\hat{C}-\kappa\right]\varepsilon_{\omega,t}\\
&\qquad + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\widehat{D}-\tilde{\beta}\varrho_\pi\right]\Big[\hat{A}\left[\rho_\zeta\zeta_{t-2}+\varepsilon_{\zeta,t-1}+\theta_\zeta\varepsilon_{\zeta,t-2}\right]+\hat{B}\varepsilon_{\zeta,t-1}\\
&\qquad + \hat{C}\varepsilon_{\omega,t-1}+\widehat{D}\left[\hat{A}\zeta_{t-2}+\hat{B}\varepsilon_{\zeta,t-2}+\hat{C}\varepsilon_{\omega,t-2}+\widehat{D}\pi_{t-3}\right]\Big]\Big].
\end{aligned}$$

Hence:

$$\mathbb{E}x_t\varepsilon_{\zeta,t} = \sigma_\zeta^2\kappa^{-1}\Big[\left[1-\tilde{\beta}(1-\varrho_\pi)\left(\widehat{D}+\rho_\zeta+\theta_\zeta\right)\right]\hat{A}+\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\hat{B}\Big],$$

$$\begin{aligned}
\mathbb{E}x_t\varepsilon_{\zeta,t-1} &= \sigma_\zeta^2\kappa^{-1}\Big[\left[1-\tilde{\beta}(1-\varrho_\pi)\left(\widehat{D}+\rho_\zeta\right)\right]\hat{A}\left(\rho_\zeta+\theta_\zeta\right)\\
&\qquad + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\widehat{D}-\tilde{\beta}\varrho_\pi\right]\left(\hat{A}+\hat{B}\right)\Big],
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}x_t\varepsilon_{\zeta,t-2} &= \sigma_\zeta^2\kappa^{-1}\Big[\left[1-\tilde{\beta}(1-\varrho_\pi)\left(\widehat{D}+\rho_\zeta\right)\right]\hat{A}\rho_\zeta\left(\rho_\zeta+\theta_\zeta\right)\\
&\qquad + \left[\left[1-\tilde{\beta}(1-\varrho_\pi)\widehat{D}\right]\widehat{D}-\tilde{\beta}\varrho_\pi\right]\left[\hat{A}\left(\rho_\zeta+\theta_\zeta\right)+\widehat{D}\left(\hat{A}+\hat{B}\right)\right]\Big].
\end{aligned}$$

Now denote by $\mathcal{T}$ the map taking the vector:

$$\widehat{m} := \begin{bmatrix} \widehat{m}_0 \\ \widehat{m}_1 \\ \widehat{m}_2 \end{bmatrix}$$

to the vector:

$$\mathcal{T}(\widehat{m}) := \begin{bmatrix} \mathbb{E} x_t \varepsilon_{\zeta,t} \\ \mathbb{E} x_t \varepsilon_{\zeta,t-1} \\ \mathbb{E} x_t \varepsilon_{\zeta,t-2} \end{bmatrix}.$$

Stochastic approximation theory relates the stability of our nonlinear difference equation to the stability of the ODE:

$$\frac{d\widehat{m}(\tau)}{d\tau} = \mathcal{T}\big(\widehat{m}(\tau)\big) - \widehat{m}(\tau).$$

The $\mathcal{T}$ map here plays the role usually played by the mapping from the perceived law of motion to the actual law of motion in the reduced form learning literature (Evans & Honkapohja 2001).

We conjecture that:

$$m := \begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix}$$

is a locally asymptotically stable point of this ODE. To check this, note that tedious algebra gives that:

$$\left.\frac{\partial \mathcal{T}(\widehat{m})}{\partial \widehat{m}}\right|_{\widehat{m}=m} = \frac{\phi_x}{\kappa \phi_\pi} \begin{bmatrix} 1 & \phi_\pi^{-1} - \tilde{\beta}(1-\varrho_\pi) & \dfrac{\phi_\pi^{-1} - \tilde{\beta}(1-\varrho_\pi)}{\phi_\pi - \rho_\zeta} \\ -\tilde{\beta}\varrho_\pi & 1 - \phi_\pi^{-1}\tilde{\beta}\varrho_\pi & \dfrac{\phi_\pi\left[\phi_\pi^{-1} - \tilde{\beta}(1-\varrho_\pi)\right] - \phi_\pi^{-1}\tilde{\beta}\varrho_\pi}{\phi_\pi - \rho_\zeta} \\ 0 & -\tilde{\beta}\varrho_\pi & \dfrac{\phi_\pi\left[1 - \tilde{\beta}(1-\varrho_\pi)\rho_\zeta\right] - \tilde{\beta}\varrho_\pi}{\phi_\pi - \rho_\zeta} \end{bmatrix}.$$

For simplicity, we assume $\phi_x \geq 0$, $\phi_\pi \geq 0$, $\kappa \geq 0$, $\tilde{\beta} \geq 0$, $\varrho_\pi \in [0,1)$, $\rho_\zeta \in [0,1)$ and $\phi_\pi \geq \left[\tilde{\beta}(1-\varrho_\pi)\right]^{-1}$. Under these assumptions, the off-diagonal elements of this matrix are all non-positive. Other cases may also go through, but for the sake of brevity we concentrate on this most relevant case. Given these assumptions, applying the Gershgorin circle theorem to the columns of this matrix gives the following upper bound on the real part of the eigenvalues of $\left.\frac{\partial \mathcal{T}(\widehat{m})}{\partial \widehat{m}}\right|_{\widehat{m}=m}$:

$$\frac{\phi_x}{\kappa\phi_\pi}\max\left\{\begin{matrix} 1+\tilde{\beta}\varrho_\pi, \phi_\pi^{-1}\left[\tilde{\beta}(\phi_\pi-\varrho_\pi)+\phi_\pi-1\right], \\ \frac{(1-\phi_\pi^{-1})(\phi_\pi-\tilde{\beta}\varrho_\pi)+\tilde{\beta}(1-\varrho_\pi)\left[1+\phi_\pi(1-\rho_\zeta)\right]-\phi_\pi^{-1}}{\phi_\pi-\rho_\zeta}\end{matrix}\right\}.$$

The first and second arguments in curly brackets here are both less than $1+\tilde{\beta}$. Taking the derivative of the third argument in curly brackets with respect to ρ_ζ produces an expression whose sign is not a function of ρ_ζ. Thus, the third argument in curly brackets is maximized at either $\rho_\zeta=0$ or $\rho_\zeta=1$. In the former case, the argument is less or equal to $1+\tilde{\beta}$ providing $\tilde{\beta}\leq 1$. In the latter case, the argument is less or equal to $1+\tilde{\beta}$ providing that $2(1-\varrho_\pi)\leq\phi_\pi$. Therefore, if $\phi_x\geq 0$, $\phi_\pi\geq 0$, $\kappa\geq 0$, $\tilde{\beta}\in[0,1]$, $\varrho_\pi\in[0,1)$, $\rho_\zeta\in[0,1)$ and:

$$\phi_\pi>\max\left\{\frac{1}{\tilde{\beta}(1-\varrho_\pi)}, 2(1-\varrho_\pi), \frac{\phi_x(1+\tilde{\beta})}{\kappa}\right\},$$

then all of the eigenvalues of $\left.\frac{\partial\mathcal{T}(\widehat{m})}{\partial\widehat{m}}\right|_{\widehat{m}=m}$ are less than one. Consequently, in this case the ODE is locally asymptotically stable, so the stochastic approximation results of Evans & Honkapohja (2001) apply. In particular, if we suppose that $\widehat{m}_0$, $\widehat{m}_1$ and $\widehat{m}_2$ are constrained to remain within a sufficiently small ball around m_0, m_1 and m_2, then the central bank's estimates of the Phillips curve parameters will converge to their true values, and the model's dynamics will converge to the determinate ones under rational expectations.

A.4 Responding to other endogenous variables in a general model

Now, suppose the central bank uses the rule:

$$i_t=r_t+\phi_\pi\pi_t+\iota\phi_z^\top z_t+\phi_\nu^\top\nu_t.$$

Here, z_t is a vector of other endogenous variables, with $z_{t,1}=r_t$, $\iota>0$ is a scalar governing the strength of response to all of them, and ν_t is an arbitrary exogenous stochastic process (potentially vector valued). As usual, we assume $\phi_\pi>1$.

Without loss of generality, we suppose that the other endogenous variables satisfy the general linear expectational difference equation:

$$0=A\mathbb{E}_t z_{t+1}+Bz_t+Cz_{t-1}+d\pi_t+E\nu_t,$$

where the coefficient matrices are such that there is a unique matrix F with eigenvalues in the unit circle such that $F = -(AF + B)^{-1}C$.[39] This condition on F just states that there is no real indeterminacy in the model. Once inflation is determined, so too is z_t. Having the same shock process entering both the monetary rule and the model's other equations is without loss of generality as it is multiplied by $\phi_\nu^\top$ and E respectively.

Now define:

$$G := -A(AF + B)^{-1}.$$

Let L be the lag operator, then note that:

$$(I - GL^{-1})(AF + B)(I - FL) = AL^{-1} + B + CL.$$

Thus, by the model's real determinacy, all of G's eigenvalues must also be inside the unit circle.

In terms of the lag operator, the model to be solved is then:

$$\mathbb{E}_t(1 - \phi_\pi^{-1}L^{-1})\pi_t = -\iota\phi_\pi^{-1}\phi_z^\top z_t - \phi_\pi^{-1}\phi_\nu^\top \nu_t,$$

$$\mathbb{E}_t(I - GL^{-1})(AF + B)(I - FL)z_t = -d\pi_t - E\nu_t.$$

Note for future reference that since ϕ_π^{-1}, G and F all have all their eigenvalues in the unit circle, $(1 - \phi_\pi^{-1}L^{-1})$, $(I - GL^{-1})$ and $(I - FL)$ are all invertible.

We conjecture a series solution of the form:

$$\pi_t = \sum_{k=0}^{\infty} \iota^k \pi_t^{(k)}, \qquad z_t = \sum_{k=0}^{\infty} \iota^k z_t^{(k)}.$$

Matching terms gives that $\pi_t^{(0)}$ solves:

$$\mathbb{E}_t(1 - \phi_\pi^{-1}L^{-1})\pi_t^{(0)} = -\phi_\pi^{-1}\phi_\nu^\top \nu_t,$$

implying that $\pi_t^{(0)}$ is determinate with:

$$\pi_t^{(0)} = -\mathbb{E}_t(1 - \phi_\pi^{-1}L^{-1})^{-1}\phi_\pi^{-1}\phi_\nu^\top \nu_t.$$

Similarly, from matching terms in the law of motion for z_t, we have that:

$$\mathbb{E}_t(I - GL^{-1})(AF + B)(I - FL)z_t^{(0)} = -d\pi_t^{(0)} - E\nu_t$$

[39] The lack of terms in $\mathbb{E}_t\pi_{t+1}$ and π_{t-1} is without loss of generality, as such responses can be included by adding an auxiliary variable $z_{t,j}$ with an equation of the form $z_{t,j} = \pi_t$.

so $z_t^{(0)}$ is also determinate (by our assumption on A, B and C) with:

$$z_t^{(0)} = -(I - FL)^{-1}(AF + B)^{-1}\mathbb{E}_t(I - GL^{-1})^{-1}\big(d\pi_t^{(0)} - E\nu_t\big).$$

Note that $\pi_t^{(0)}$ can be treated as exogenous for solving for $z_t^{(0)}$, as the causation only runs one way, from $\pi_t^{(0)}$ to $z_t^{(0)}$.

Now suppose that we have established that $\pi_t^{(k)}$ and $z_t^{(k)}$ are determinate for some $k \in \mathbb{N}$, with a determined solution not a function of higher order terms. (We have already proven the base case of $k = 0$.) We seek to prove that $\pi_t^{(k+1)}$ and $z_t^{(k+1)}$ are also determinate. Matching terms again gives that:

$$\mathbb{E}_t(1 - \phi_\pi^{-1}L^{-1})\pi_t^{(k+1)} = -\phi_\pi^{-1}\phi_z^\top z_t^{(k)},$$

so $\pi_t^{(k+1)}$ is also determinate, with:

$$\pi_t^{(k+1)} = -\mathbb{E}_t(1 - \phi_\pi^{-1}L^{-1})^{-1}\phi_\pi^{-1}\phi_z^\top z_t^{(k)},$$

where we used the inductive hypothesis that $z_t^{(k)}$ is already determined, and so it is effectively exogenous for the purpose of determining $\pi_t^{(k+1)}$. Then from matching terms in the law of motion for z_t:

$$\mathbb{E}_t(I - GL^{-1})(AF + B)(I - FL)z_t^{(k+1)} = -d\pi_t^{(k+1)},$$

so $z_t^{(k+1)}$ is also determinate, with:

$$z_t^{(k+1)} = -(I - FL)^{-1}(AF + B)^{-1}\mathbb{E}_t(I - GL^{-1})^{-1}d\pi_t^{(k+1)},$$

much as before. This completes our proof by induction, establishing that there is a series solution of the given form.

The only remaining thing to check is that the series does indeed converge for sufficiently small ι. This follows immediately from the product structure of the solution above, which means that the variances of $z_t^{(k)}$ and $\pi_t^{(k)}$ must be $O\big(h^k\big)$ for some $h \geq 1$. Hence for sufficiently small ι, the model is determinate. I.e., given the Taylor principle is satisfied, a sufficiently small response to other endogenous variables will not break determinacy.

A.5 If inflation is identical, other endogenous variables are identical

Let x_t and $\tilde{x}_t$ be vectors stacking the endogenous variables other than inflation in the economy with our rule and the economy with the alternative rule, respectively. We assume without loss of generality that they are all zero in steady state. By linearity, the equations other than the monetary rule or monetary policy first order condition must have the form:

$$0 = \mathbb{A}x_{t-1} + \mathbb{a}\pi_{t-1} + \mathbb{B}x_t + \mathbb{b}\pi_t + \mathbb{C}\mathbb{E}x_{t+1} + \mathbb{c}\mathbb{E}\pi_{t+1} + \sum_{n=1}^{N} \mathbb{d}_n \varepsilon_{n,t}\,, \tag{10}$$

in the economy with our rule, and they must have the form:

$$0 = \mathbb{A}\tilde{x}_{t-1} + \mathbb{a}\tilde{\pi}_{t-1} + \mathbb{B}\tilde{x}_t + \mathbb{b}\tilde{\pi}_t + \mathbb{C}\mathbb{E}\tilde{x}_{t+1} + \mathbb{c}\mathbb{E}\tilde{\pi}_{t+1} + \sum_{n=1}^{N} \mathbb{d}_n \varepsilon_{n,t},$$

in the economy with the alternative rule. (Here, $\mathbb{A}$, $\mathbb{B}$ and $\mathbb{C}$ are square matrices, while $\mathbb{a}$, $\mathbb{b}$ and $\mathbb{c}$ are scalars, and $\mathbb{d}_1, \dots, \mathbb{d}_N$ are vectors.) Since $\pi_t \equiv \tilde{\pi}_t$, $x_t \equiv \tilde{x}_t$ must solve equation (9). It will be the unique solution providing the model has no source of indeterminacy other than perhaps monetary policy. For example, in a three equation NK model, given that $\pi_t \equiv \tilde{\pi}_t$, the Phillips curve implies that the output gap must agree in the two economies, thus the Euler equation then implies that the interest rate must also agree.

A.6 Solution properties of first welfare example

Recall, that for $k > 1$ the solution must satisfy the recurrence relation:

$$\theta_k + \frac{\lambda}{\kappa^2}\left(\theta_k - \beta\theta_{k+1}\right) - \beta\frac{\lambda}{\kappa^2}\left(\theta_{k-1} - \beta\theta_k\right) = 0.$$

The characteristic equation of this recurrence relationship has roots:

$$\frac{\left(1 + \frac{\lambda}{\kappa^2} + \beta^2\frac{\lambda}{\kappa^2}\right) \pm \sqrt{\left(1 + \frac{\lambda}{\kappa^2} + \beta^2\frac{\lambda}{\kappa^2}\right)^2 - \left(2\beta\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}$$
$$= \frac{\left(1 + \frac{\lambda}{\kappa^2} + \beta^2\frac{\lambda}{\kappa^2}\right) \pm \sqrt{\left(1 + (1+\beta)^2\frac{\lambda}{\kappa^2}\right)\left(1 + (1-\beta)^2\frac{\lambda}{\kappa^2}\right)}}{2\beta\frac{\lambda}{\kappa^2}}.$$

The positive root satisfies:

$$\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)+\sqrt{\left(1+(1+\beta)^2\frac{\lambda}{\kappa^2}\right)\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)}}{2\beta\frac{\lambda}{\kappa^2}}$$

$$>\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)+\sqrt{\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)}}{2\beta\frac{\lambda}{\kappa^2}}$$

$$=\frac{1+\frac{\lambda}{\kappa^2}-\beta(1-\beta)\frac{\lambda}{\kappa^2}}{\beta\frac{\lambda}{\kappa^2}}>\frac{1+\frac{\lambda}{\kappa^2}-(1-\beta)\frac{\lambda}{\kappa^2}}{\beta\frac{\lambda}{\kappa^2}}=1+\frac{1}{\beta\frac{\lambda}{\kappa^2}}>1.$$

The negative root satisfies:

$$\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)^2-\left(2\beta\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}$$

$$>\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}=0,$$

and:

$$\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+(1+\beta)^2\frac{\lambda}{\kappa^2}\right)\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)}}{2\beta\frac{\lambda}{\kappa^2}}$$

$$<\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)\left(1+(1-\beta)^2\frac{\lambda}{\kappa^2}\right)}}{2\beta\frac{\lambda}{\kappa^2}}=1.$$

Hence, the positive root is greater than 1, while the negative root is in $(0,1)$. Thus for $k \geq 1$:

$$\theta_k=\theta_1\left[\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta^2\frac{\lambda}{\kappa^2}\right)^2-\left(2\beta\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}\right]^{k-1}.$$

Hence, θ_0, θ_1 and θ_2 are the unique solution of the three linear (in θ_0, θ_1 and θ_2) equations:

$$\theta_0+\frac{\lambda}{\kappa^2}(\theta_0-\beta\theta_1-1)=0,$$

$$\theta_1 + \frac{\lambda}{\kappa^2}(\theta_1 - \beta\theta_2) - \beta\frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\theta_2 = \theta_1 \left[\frac{\left(1 + \frac{\lambda}{\kappa^2} + \beta^2 \frac{\lambda}{\kappa^2}\right) - \sqrt{\left(1 + \frac{\lambda}{\kappa^2} + \beta^2 \frac{\lambda}{\kappa^2}\right)^2 - \left(2\beta\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}} \right].$$

A.7 Solution under discretion of first welfare example

Under discretion, we have the standard first order condition:

$$\pi_t + \frac{\lambda}{\kappa}x_t = 0,$$

i.e.:

$$\kappa \sum_{k=0}^{\infty} \theta_k \omega_{t-k} + \frac{\lambda}{\kappa} \sum_{k=0}^{\infty} (\theta_k - \beta\theta_{k+1} - \mathbb{1}[k=0])\omega_{t-k} = 0,$$

so:

$$\theta_0 + \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\forall k \geq 1, \qquad \theta_k + \frac{\lambda}{\kappa^2}(\theta_k - \beta\theta_{k+1}) = 0.$$

The latter recurrence relation has the general solution $\theta_k = \theta_1 \left(\frac{\kappa^2}{\beta\lambda} + \frac{1}{\beta}\right)^{k-1}$, which is explosive as $\beta < 1$. Thus, we must have $\theta_1 = \theta_2 = \cdots = 0$. Hence, $\theta_0 = \frac{\lambda}{\lambda+\kappa^2}$.

A.8 Solution under the timeless perspective of first welfare example

The timeless perspective (Woodford 1999) leads to the first order condition:

$$\pi_t + \frac{\lambda}{\kappa}(x_t - x_{t-1}) = 0,$$

i.e.:

$$\kappa \sum_{k=0}^{\infty} \theta_k \omega_{t-k} + \frac{\lambda}{\kappa} \left[\sum_{k=0}^{\infty} (\theta_k - \beta\theta_{k+1} - \mathbb{1}[k=0])\omega_{t-k} - \sum_{k=1}^{\infty} (\theta_{k-1} - \beta\theta_k - \mathbb{1}[k-1=0])\omega_{t-k} \right] = 0,$$

so:

$$\theta_0 + \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\theta_1 + \frac{\lambda}{\kappa^2}(\theta_1 - \beta\theta_2) - \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\forall k > 1, \qquad \theta_k + \frac{\lambda}{\kappa^2}(\theta_k - \beta\theta_{k+1}) - \frac{\lambda}{\kappa^2}(\theta_{k-1} - \beta\theta_k) = 0.$$

The roots of the characteristic equation corresponding to the latter recurrence relation are:

$$\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)\pm\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)^2-4\beta\left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}.$$

The positive root satisfies:

$$\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)+\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)^2-4\beta\left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}>\frac{\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}}{2\beta\frac{\lambda}{\kappa^2}}=\frac{1+\beta}{2\beta}>1.$$

The negative root satisfies:

$$\begin{aligned}&\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)^2-4\beta\left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}\\&\qquad>\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}=0,\end{aligned}$$

and:

$$\begin{aligned}&\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)^2-4\beta\left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}\\&\qquad=\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{1+(1-\beta)^2\left(\frac{\lambda}{\kappa^2}\right)^2+2(1+\beta)\frac{\lambda}{\kappa^2}}}{2\beta\frac{\lambda}{\kappa^2}}\\&\qquad<\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{1+(1-\beta)^2\left(\frac{\lambda}{\kappa^2}\right)^2+2(1-\beta)\frac{\lambda}{\kappa^2}}}{2\beta\frac{\lambda}{\kappa^2}}\\&\qquad=\frac{\left(1+\frac{\lambda}{\kappa^2}+\beta\frac{\lambda}{\kappa^2}\right)-\sqrt{\left(1+(1-\beta)\frac{\lambda}{\kappa^2}\right)^2}}{2\beta\frac{\lambda}{\kappa^2}}=\frac{2\beta\frac{\lambda}{\kappa^2}}{2\beta\frac{\lambda}{\kappa^2}}=1.\end{aligned}$$

Hence, the positive root is greater than 1, while the negative root is in $(0,1)$. Thus for $k \geq 1$:

$$\theta_k = \theta_1 \left[\frac{\left(1 + \frac{\lambda}{\kappa^2} + \beta \frac{\lambda}{\kappa^2}\right) - \sqrt{\left(1 + \frac{\lambda}{\kappa^2} + \beta \frac{\lambda}{\kappa^2}\right)^2 - 4\beta \left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta \frac{\lambda}{\kappa^2}} \right]^{k-1}.$$

Hence, θ_0, θ_1 and θ_2 are the unique solution of the three linear (in θ_0, θ_1 and θ_2) equations:

$$\theta_0 + \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\theta_1 + \frac{\lambda}{\kappa^2}(\theta_1 - \beta\theta_2) - \frac{\lambda}{\kappa^2}(\theta_0 - \beta\theta_1 - 1) = 0,$$

$$\theta_2 = \theta_1 \left[\frac{\left(1 + \frac{\lambda}{\kappa^2} + \beta \frac{\lambda}{\kappa^2}\right) - \sqrt{\left(1 + \frac{\lambda}{\kappa^2} + \beta \frac{\lambda}{\kappa^2}\right)^2 - 4\beta \left(\frac{\lambda}{\kappa^2}\right)^2}}{2\beta \frac{\lambda}{\kappa^2}} \right].$$