European Central Bank

EUROSYSTEM

Working Paper Series

Frantisek Masek, Jan Zemlicka

Average inflation targeting:
how far to look into the past and the
future?

No 2955

Abstract

We analyze the optimal window length in the average inflation targeting rule within a Behavioral THANK model. The central bank faces an occasionally binding effective lower bound (ELB) or persistent supply shocks, and can also use quantitative easing. We show that the optimal averaging period is infinitely long given a conventional degree of myopia. Finite yet long-lasting windows dominate for higher cognitive discounting; i.e., the makeup property is shown to be qualitatively resistant to deviation from rational expectations. We point out that the optimal window may depend on the speed of return to the target path. We solve the model both locally and globally to disentangle the effects of uncertainty due to the ELB. The welfare loss difference between solution techniques is considerably decreasing in the degree of history dependence.

Non-technical summary

The change in monetary policy strategy that the US Federal Reserve made in August 2020 attracted a great deal of interest to the average inflation targeting (AIT) rule (see Fed, 2020). The alteration of monetary policy strategy was the outcome of a long-term investigation of makeup monetary policy rules in response to an environment of the low natural rate of interest featured with a high likelihood of the effective lower bound (ELB) situation. Although price level targeting (PLT) was most studied early on, AIT later came under more scrutiny, given that it is seen by many as a middle ground between inflation targeting (IT) and PLT.

An intriguing characteristic of the Fed's new framework is the absence of an exactly specified averaging window length for inflation. In this article, we analyze the welfare consequences of varying the degree of history-dependence in a model that behaves realistically in terms of monetary policy. Further, we analyze the impact of the speed of return of average inflation to its target path. While the window length has been studied in other articles (see Budianto et al., 2023; Amano et al., 2020; Coulter et al., 2022), to the best of our knowledge, we are the first to study the interplay of the averaging window and the speed of return to the target.

Our analysis utilizes a New Keynesian model which resolves two problematic aspects of standard Rational Agent New Keynesian (RANK) models. We use the tractable heterogeneity of households from Bilbiie (2024) and cognitive discounting of Gabaix (2020). Hence, our model can resolve the forward guidance puzzle while keeping the amplification of the contemporary monetary policy shock consistent with the findings stemming from the Heterogeneous Agents New Keynesian (HANK) literature. The model behaves realistically regarding both future and contemporary monetary policy. Resolution of the forward guidance puzzle is especially crucial to the validity of our analysis.

Considering the natural rate of interest shock and the lower bound on the policy rate, the optimal averaging period is infinitely long (i.e., the PLT is optimal) if we impose only moderate cognitive discounting. To make the ELB situation less severe, we extend the benchmark analysis by allowing the central bank to use quantitative easing (QE) when the economy is at the lower bound. The QE does not alter the core results,

though. We obtain the same results when we include past inflation outcomes in the expectations formation, and deviate even more from the rational expectations equilibrium.

Taking into account the evidence of Coibion et al. (2023) of US households' flawed understanding of the Fed's announcement of AIT, we study the effects of stronger cognitive discounting. We demonstrate that the degree of myopia in expectations formation alters the welfare comparison of AIT and IT only quantitatively. A higher degree of myopia attenuates the superiority of AIT over IT. However, we show that, so long as the central bank does not try to close the gap between average inflation and the target path too quickly, IT does not generate a lower welfare loss than AIT. The welfare loss difference between AIT and IT diminishes substantially, but AIT remains superior.

Our results differ from Budianto et al. (2023), where stronger degrees of cognitive discounting result in the welfare superiority of shorter window lengths closer to IT. The difference arises due to distinct monetary policy rules. In Budianto et al. (2023), the central bank conducts monetary policy under optimal discretion. We work with a feedback rule and calibrate the response parameter to average inflation to give the central bank more time to return to the target path. In contrast, in Budianto et al. (2023), the central bank chooses inflation consistent with the target path immediately, given the optimal discretion setting. In comparison to Budianto et al. (2023), we highlight the importance of the assumption about the speed of return to the target path when analyzing makeup rules.

To separate a downward inflation bias resulting from the presence of the ELB, we solve the model both locally and globally. The difference between the local and global solutions accounts for the effects of uncertainty about hitting the ELB in the future. The welfare loss difference between the solution techniques is considerably decreasing in the degree of history dependence: AIT helps to mitigate the downward inflation bias. However, the difference starts to disappear under stronger cognitive discounting as exogenous risk is discounted more.

1 Introduction

The falling natural rate of interest observed in recent decades has directed attention to alternative monetary policy rules that may be better equipped for a world with the effective lower bound (ELB) on the policy rate. These are so-called makeup rules, such as price level targeting (PLT) and average inflation targeting (AIT).[1][2] AIT has even become a new monetary policy strategy of the US Federal Reserve (see Fed, 2020).[3]

As the Federal Reserve enters a new monetary policy review phase, we aim to shed light on two key components of AIT that have not yet been fully addressed. First, we study the optimal window length of AIT. The Federal Reserve has not provided any explicit period over which it intends to average inflation. Utilizing a New Keynesian model featuring a bounded rationality extension and heterogeneity on the households side, we show that the welfare-optimal averaging period is infinite (equivalent to PLT) so long as the deviation from rational expectations is not substantial.[4] Interestingly, even if we deviate further from rational expectations (stronger cognitive discounting of Gabaix (2020)), the optimal window length is still long-lasting and closer to PLT than IT. We show that makeup rules may be relatively resistant to deviation from rational expectations formation. The superiority of PLT and AIT over IT shrinks but it does not completely disappear, even under severe degrees of myopia. In this regard, our results differ from the recent analysis of AIT in Budianto et al. (2023), who show that stronger cognitive discounting leads to a shorter optimal window length.

The second characteristic of makeup rules that we study is the effects of the speed of return to the target path. We offer an explanation of the discrepancy between Bu-

---

[1]To mention the literature concerning ELB, we point to Krugman (1998); Eggertson and Woodford (2003); Svensson (2001); Bernanke (2000); Adam and Billi (2006, 2007) and Nakov (2008), and to more recent Mertens and Williams (2019) and Svensson (2020).

[2]There has recently been a vivid debate about the direction of the natural rate of interest (see Schnabel, 2024). Some point to a possible increase, which would lessen the problem of ELB (see Benigno et al., 2024). However, many others have remained sceptical that the period of low interest rate is beyond us, meaning the ELB remains relevant (see Obstfeld, 2023; Bäcker-Peral et al., 2024).

[3]Clarida (2020) clarifies a wide range of aspects of the new framework.

[4]Firstly, the model delivers amplification of a contemporary monetary policy change through an indirect general equilibrium effect given by the presence of heterogeneous agents in line with Bilbiie (2024). Second, it can rule out the forward guidance puzzle due to the cognitive discounting of Gabaix (2020); i.e., overly strong sensitivity of current variables to the expected path of the real interest rate.

dianto et al. (2023) and our results based on the difference in the monetary policy reaction. While Budianto et al. (2023) allow the central bank to operate under optimal discretion, we work with a standard feedback rule of the Taylor (1993) type. We stress that such a difference has a profound effect on the speed of return of average inflation to the target path. The return is assumed to be immediate under optimal discretion, while the Taylor rule enables variations of it by changing the value of the elasticity parameter. We show that our results become consistent with Budianto et al. (2023) only under unrealistically high values of the Taylor rule parameter; only when the central bank wants to close the gap between average inflation and its target path extremely quickly. To the best of our knowledge, we are the first to discuss the interaction of the averaging period and the speed of return to the target path.

Additionally, we rule out the explanation based on deterioration of the expectations channel when cognitive discounting is strong and the central bank wants to move average inflation to the target path quickly. One might assume that this invokes sharper reactions of the policy rate (because the real interest rate does not move sufficiently due to more severe myopia), resulting in a higher variance of output and inflation. However, we document that this intuitive explanation does not fit our results nor those of Budianto et al. (2023). It is not primarily worse-performing makeup rules under strong cognitive discounting and quick speed of return that enables shorter window lengths closer to IT to perform in a superior manner. The cause of the IT superiority is that IT with a very swift policy rate reaction (high value of Taylor rule coefficient) works very efficiently with strong myopia in the ELB situation.

Our results show that makeup rules may be a desirable policy tool in the ELB situation even when agents deviate further from rational expectations formation. We also stress that the speed of return does not substantially alter welfare if the central bank imposes a long-lasting AIT or even PLT. Nevertheless, we point out that more insights need to be provided on this characteristic of makeup rules, given that, so far, the discussion has been mainly about the length of the averaging period and not about the speed of return to the target path. We also show that a stronger reaction to deviations of inflation in IT may offer another way to handle ELB without necessarily going in the direction of makeup rules.

In addition to our main results, we conduct various extensions. We show that the

main results remain unchanged when we incorporate a backward-looking component into expectations formation, or if we enable the central bank to also use unconventional monetary policy (QE) once the economy hits ELB. None of the alterations of the baseline model alter the core results in a qualitative way.

As an additional result, we quantify the magnitude of the so-called deflationary (or downward inflation) bias, which has played an important role in the ELB literature (see Eggertsson, 2006; Penalver and Siena, 2024). To separate the bias due to the presence of the ELB, we solve the model both locally and globally. The difference between the two solutions accounts for the effects of uncertainty about hitting the ELB in the future. The welfare loss difference between solution techniques is considerably decreasing in the degree of history dependence. Put differently, AIT helps to mitigate downward inflation bias. However, the difference starts to disappear under stronger cognitive discounting as the uncertainty is discounted more.

Related literature. A pioneering analysis of AIT within an NK framework is found in Nessén and Vestin (2005). The authors show that, in a purely forward looking framework, PLT dominates AIT. However, when both backward- and forward-looking components are mixed within the Phillips curve, AIT may be superior to both IT and PLT.

Budianto et al. (2023) study the welfare optimal averaging window length for inflation. The authors work with the Behavioral New Keynesian model of Gabaix (2020) to attenuate the strength of the expectations channel of monetary policy. The results of their analysis show that, so long as the cognitive discounting parameter is not too small (i.e., cognitive limitations are not too high), AIT performs better than IT and increases agents' welfare. The resulting length of the optimal averaging window is infinitely long. Nevertheless, when the level of myopia is higher, the optimal averaging period becomes finite and gains resulting from switching to AIT are much smaller. Dobrew et al. (2023) also utilize Gabaix (2020) and conclude that makeup rules lose their advantage over IT under stronger myopia. However, they highlight that an exponential moving average (MA) for AIT performs substantially better than an arithmetic MA.

Feiveson et al. (2020) analyze the behavior of a HANK model under AIT and PLT. They show that history-dependent strategies can help alleviate the adverse effects of ELB on unemployment and inflation. Beyond business cycle fluctuations, they also

discuss distributional issues. Arias et al. (2020) and Hebden et al. (2020) also study the makeup policy regimes within the Fed's review process, as does Feiveson et al. (2020). Arias et al. (2020) find that history-dependent monetary policy regimes may be more beneficial than IT. However, the gains are moderate, and the authors note that there are issues that can impact practical implementation of these strategies. Hebden et al. (2020) investigate how robust makeup rules are to changes in inflation expectations formation. They conclude that history-dependent strategies may be effective even when a substantial fraction of the general public is uninformed about the monetary policy rule.

The AIT rule is also studied in a HANK model by Djeutem et al. (2022). The authors show that history-dependent rules are superior in their modelling framework. IT and AIT can potentially dominate an entirely history-dependent PLT only when the central bank is concerned about inequality in its loss function. Jia and Wu (2021) show that the absence of an exact window length in the Fed's new monetary framework might be intentional, and could be beneficial from the central bank's perspective. In contrast, Honkapohja and McClung (2021) show that the use of AIT can pose significant macroeconomic instability compared to IT or different makeup rules such as PLT within a learning type of model.

Coulter et al. (2022) analyze the effect of the change in the Fed's monetary policy framework on the inflation spike in subsequent years. Deploying both a quasi-experimental approach using the synthetic control method and a structural analysis utilizing the Martínez-García (2021) New Keynesian model, they show that the switch from IT to AIT in August 2020 can explain only a minor part of the inflation surge in the subsequent years. Piergallini (2022) shows that a high weight on the distant past under AIT ensures local determinacy and eliminates the liquidity trap situation.

Outline. The rest of the paper is organized as follows. Section 2 describes the model we use for the analysis. Section 3 presents our core results, and section 4 expands them to the case of a backward-looking component in the expectations formation, unconventional monetary policy, and the role of supply shocks. Section 5 concludes.

2 Model

We use a New Keynesian model in discrete time extended by heterogeneous and bounded rationality elements to deliver a more realistic environment for investigating makeup monetary policy rules. The demand side consists of a tractable version of the HANK (THANK) model by Bilbiie (2024), and bounded rationality defined by the approach used by Gabaix (2020) in the form of cognitive discounting. The myopic expectation is applied to both output and inflation; i.e., it also affects the real interest rate.[5]. Such a framework can generate both features of monetary policy that a fully rational New Keynesian (RANK) model cannot cope with, which have recently been the subject of much discussion. These are the monetary policy amplification through indirect general equilibrium effect, and overly strong sensitivity on the expected path of the real interest rate-forward guidance puzzle.[6] Incorporating the heterogeneous and behavioral features from above can resolve both issues at once. Thus, following Bilbiie (2024), there is no *Catch-22* defined by the trade-off in terms of the ability to solve always only one of the described characteristics within the HANK framework.[7]

We derive the full model in our online appendix (link here). The dynamic optimization problem of households leads to a Behavioral Dynamic IS Curve:

$$\tilde{y}_t = \psi_f \, \mathbb{E}_t \, \tilde{y}_{t+1} - \psi_c \frac{1}{\sigma}(i_t - \overline{m} \, \mathbb{E}_t \, \pi_{t+1} - r_t^{nat}), \tag{1}$$

where $\psi_f = \overline{m}\delta = \overline{m}\big[1 + (\chi - 1)\frac{1-s}{1-\lambda\chi}\big]$ and $\psi_c = \frac{(1-\lambda)}{(1-\lambda\chi)}$.

To clarify, $\overline{m}$ denotes the cognitive discounting from the Behavioral New Keynesian model of Gabaix (2020). The probability of staying a saver household type in a two-state switching Markov process from Bilbiie (2024) in which households can switch between being savers and hand-to-mouth types is s. The IS curve is approximated around the ergodic distribution λ denoting the constant share of hand-to-mouth households given that we focus on stationary equilibria. The key parameter $\chi = 1 + \varphi(1 - \frac{\zeta^D}{\lambda})$ with ζ^D defining a fiscal redistribution from savers - firm owners - towards hand-to-mouth

---

[5]Thus, the household side works with a setup that is similar to that of Pfäuti and Seyrich (2023)

[6]We discuss in detail the suitability of the model for analysis of makeup rules in the appendix.

[7]Note that the trade-off may eventually be resolved without extending the HANK structure by the bounded rationality component. It is necessary to combine cyclicality of inequality and income risk such that they have opposite signs, as discussed in Bilbiie (2024).

households drives the cyclicality of the consumption inequality in the model. We assume that $\chi > 1$, resulting in the countercyclical inequality ([Bilbiie, 2024](#)). The natural rate of interest follows an AR(1) process.

The supply side of the model is based on the approach of [Gabaix](#) ([2020](#)); i.e., intermediate firms incorporate the cognitive discounting mechanism. The discounting takes into account the degree of price stickiness; a higher parameter of price rigidity leads to a more forward-looking New Keynesian Phillips curve. We allow for more cognitive discounting parameters, as, in addition to the general cognitive discounting parameter, the firms are also myopic with respect to future inflation and the path of marginal costs. We use simple constant returns to scale production function $Y_t = N_t$.[8] The resulting Behavioral Phillips Curve is as follows:

$$\pi_t = \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa \tilde{y}_t, \tag{2}$$

where $M^f = \overline{m}\left[\theta + \frac{1-\beta\theta}{1-\beta\theta\overline{m}} m_\pi^f (1-\theta)\right]$. Moreover, $\kappa = m_y^f \overline{\kappa}$ and $\overline{\kappa} = \omega(\varphi + \sigma)$ while $\omega = \frac{(1-\theta)(1-\theta\beta)}{\theta}$. Parameters σ and φ are conventional inverse intertemporal elasticity of substitution and inverse Frisch elasticity from the household utility function. Simultaneously, a well-known notation is used for the subjective discount factor (β) and the probability of adjusting prices of the intermediate firms in the framework of [Calvo](#) ([1983](#)) ($1 - \theta$). The specific cognitive discounting parameters related to myopia with respect to inflation and marginal costs are m_π^f and m_y^f, respectively.

2.1 Monetary policy

Below, we formally present the monetary policy rule we use in this analysis. Unlike [Budianto et al.](#) ([2023](#)), who are focused on the situation of a central bank acting under discretion, we work with a feedback rule of the [Taylor](#) ([1993](#)) type. However, the original linear function of endogenous variables must be updated into a truncated version of the feedback rule to incorporate the lower bound.

To express the rule for average inflation targeting, we first display the definition

---

[8]We explicitly define the output gap as $\tilde{y}_t = y_t - y_t^*$, where y_t^* denotes the behavior of the natural output. However, in our model $y_t^* = 0$, as comes when we allow for all prices to be flexible, zero inflation, and constant marginal costs and markup. In other words, the fact that there is no exogenous technological shock in our production function results in $\tilde{y}_t = y_t$.

of the exponential moving average process for inflation:[9]

$$\pi_t^a = \pi_t + (1 - \xi)\pi_{t-1}^a, \tag{4}$$

where $\xi \in [0, 1]$ holds. We use the exponential moving average process to enable us to solve the model utilizing global techniques. Now we can define the average inflation targeting rule:[10]

$$i_t = \max\left\{0, \rho_i i_{t-1} + \left(1 - \rho_i\right)\left[\left(\frac{1}{\beta} - 1\right) + \phi_{\pi a}\pi_t^a + \phi_y \tilde{y}_t\right]\right\} \tag{5}$$

where one can see that this rule boils down to the inflation targeting and price level targeting regimes at the boundaries of the interval for ξ. Hence, the interior values represent the average inflation targeting rule for different window lengths.

Given the approximation properties of the exponential MA toward the arithmetic MA, one can transform the value ξ in the smooth process into k periods expression in the discrete MA as $k = \frac{2}{\xi} - 1$ for $0 < \xi \leq 1$. See Nahmias and Olsen (2015) for a thorough discussion of using exponential smoothing as the arithmetic MA approximation. [11]

2.2 Welfare function

We use a micro-founded welfare function that is consistent with Bilbiie (2024), given that the cognitive discounting of Gabaix (2020) does not necessarily enter into

---

[9]Note that expression 4 is isomorphic to the convex combination of the current and past average inflation $\pi_t^a = \xi\pi_t + (1 - \xi)\pi_{t-1}^a$ for values of $0 < \xi \leq 1$ when we adjust the elasticity for the average inflation in the monetary rule 5. Hence, the rule would be:

$$i_t = \max\left\{0, \rho_i i_{t-1} + \left(1 - \rho_i\right)\left[\left(\frac{1}{\beta} - 1\right) + \frac{\phi_{\pi a}}{\xi}\pi_t^a + \phi_y \tilde{y}_t\right]\right\} \tag{3}$$

Naturally, the rule needs to be substituted by the price level targeting $p_t = \pi_t + p_{t-1}$ when $\xi = 0$. This happens automatically in equation 4. We work with this expression straight, given it is less cumbersome in notation.

[10]Nevertheless, we abstain from smoothing in the monetary policy rules and assign $\rho_i = 0$. We do so for decreasing dimensionality as we solve the model also globally. This is the same as in Christiano et al. (2011) and Fernández-Villaverde et al. (2015) who also omit smoothing for such a reason.

[11]Coulter et al. (2022) shows a comparison of the processes within monetary policy rules, highlighting different weighting of the past periods. While in the arithmetic MA, the weight of a given period remains constant until it drops to zero at the end of the averaging period, and the exponential MA works with a decay process; i.e., the weight of a given period decreases as it falls further into the past.

welfare loss derivations.[12] In this case, consumption inequality plays a role in the welfare loss function. However, considering only shocks driving no wedge between the inequality and aggregate output gap allows us to reshuffle the welfare function into a form that is isomorphic to the benchmark welfare function consisting of inflation and output gap variance. Nonetheless, inequality increases the output gap weight compared to the benchmark welfare function stemming from the plain RANK model.

The model-consistent welfare function has the following form:[13]

$$\mathbb{W} = -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \pi_t^2 + \alpha_y \tilde{y}_t^2 + \alpha_\gamma \gamma_t^2 \right\}, \tag{6}$$

where γ_t stands for the consumption inequality. The output weight is defined conventionally as $\alpha_y = \frac{\kappa}{\epsilon}$ while the weight assigned to inequality is $\alpha_\gamma = \alpha_y \lambda (1-\lambda) \sigma \varphi^{-1}$.[14]

Further, we follow Bilbiie (2024) and consider only shocks that drive no wedge between the inequality and aggregate output gap relationship; i.e., $c_t^S - c_t^H = \frac{1-\chi}{1-\lambda} y_t$ always holds. Substituting the inequality equation into the welfare loss function gives us:

$$\mathbb{W} = -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \pi_t^2 + \alpha \tilde{y}_t^2 \right\}, \tag{7}$$

where $\alpha = \alpha_y \left[1 + \frac{\lambda}{1-\lambda} \frac{\sigma}{\varphi} (1-\chi)^2 \right]$. Put differently, the welfare loss function boils down to an isomorphic case of the RANK model. However, the inequality term increases the weight of the output gap relative to the RANK welfare loss function.

Taking into account that the value of $\mathbb{W}$ itself has no particular meaning, we follow Billi (2017) and express the results in the form of a welfare-equivalent consumption transfer. Nakata and Schmidt (2019) and Budianto et al. (2023) use an identical approach. Hence, we interpret the welfare loss in terms of a perpetual share of steady state consumption that would satisfy indifference between stationary and stochastic

---

[12]As stressed by Gabaix (2020), this approach is in much of behavioral economics; i.e., behavioral agents use heuristics in their behavior, yet they experience their utility in the same way as rational agents. Hence, objective, not subjective, expectations are used in the derivation.

[13]The derivation process is equivalent to Bilbiie (2024).

[14]Notice that, when applying cognitive discounting for the marginal cost path, the output weight in the loss function changes. However, we want to have constant α_y to compare the welfare loss for different degrees of myopia. Thus, we calibrate $m_y^f = 1$ for the purpose of α_y although we allow m_y^f to vary in the slope parameter of equation 2.

economies from the welfare perspective. Formally, we work with:

$$W = (1 - \beta) \frac{\epsilon}{\kappa} (\sigma + \varphi) \mathbb{E}(\mathbb{W}), \qquad (8)$$

where $\mathbb{W}$ is defined in 7 and the expectation is taken towards the unconditional distribution of the shock.

2.3   Baseline parameterization

We tie the baseline calibration of the model to the current NK literature related to behavioral and heterogeneous features. The model is calibrated to the U.S. economy. The complete list of the baseline calibration is in table 1 in the appendix. We briefly elaborate on the parameterization that differentiates the model from a conventional RANK model.

Considering the baseline calibration, the THANK extension of the dynamic IS curve in the form of $\frac{(1-\lambda)}{(1-\lambda\chi)}$ is equal to 1.3109. The additional compounded parameter in the IS curve is evaluated as follows: $\psi_f = \overline{m}\delta = \overline{m}\left[1 + (\chi - 1)\frac{1-s}{1-\lambda\chi}\right] = 0.8935$. Considering the PC, we get: $M^f = \overline{m}\left[\theta + \frac{1-\beta\theta}{1-\beta\theta\overline{m}}m_\pi^f(1-\theta)\right] = 0.7736$ while the inflation elasticity on output gap is $\kappa = m_y^f\overline{\kappa} = m_y^f\omega(\varphi + \sigma) = 0.0776$. Note that when assigning $\lambda = 0$ and $\overline{m} = m_y^f = m_\pi^f = 1$, we are back in the textbook three equation NK model of Galí (2015). Hence, one can envision the textbook model as a special case of the model used in this article.

Note that we keep the Taylor rule parameters constant across different window lengths ($\phi_{\pi a} = 1.5$ and $\phi_y = 0.125$). We do not want to optimize them for each averaging period on purpose. Doing so would mean changing time horizon until which the central bank wants to return to the target path. Thus, we would have different horizon for each window length; i.e., the averaging windows would not be strictly comparable.

3   Welfare optimal window length

Below we present the results of our welfare analysis.[15] Considering that the analysis is primarily about the ELB situation, we are first concerned with the demand - the

---

[15]We conduct the welfare analysis using 2,000 simulations across 1,200 periods, with 200 being discarded as burn-ins.

natural rate of interest - shock. Hence, we compute the optimal window length for AIT. Moreover, we study how our results depend on cognitive discounting, QE extension, and partly backward-looking expectations. Second, we also look for the optimal window length in the presence of supply shocks. In addition to the usual AR(1) process, we include persistence in the innovations to study the AIT in the environment of persistent mark-up shocks resembling the period in which the Federal Reserve implemented the new monetary policy rule.

As noted, we run the analysis for two different solution techniques.[16] We use the piecewise first-order perturbation solution in line with Guerrieri and Iacoviello (2015), as it may ambiguously be considered the most well-known local solution for handling OBC problems. Simultaneously, we use a global solution to compare the effect of uncertainty that the method of Guerrieri and Iacoviello (2015) abstains from. For this purpose, we use the collocation method with linear splines as basis functions.[17]

3.1 Optimal averaging period in AIT

We compute the optimal averaging window length of the AIT regime under the simple feedback rule defined earlier. We compute the welfare defined by equation 7 for a grid of values of $\xi \in [0, 1]$. Consider that two extreme cases, 0 and 1, transform average inflation targeting into price level targeting and inflation targeting, respectively.

Baseline results. In figure 1 we depict the optimal window length considering the demand shock and the ZLB. The welfare loss values stem from equation 7. The results are generated by both solution techniques. The outcome is in line with Budianto et al. (2023), who show that, in the optimization problem where a central bank conducts monetary policy under discretion, the optimal window is infinitely long within the model structure from Gabaix (2020). We reach the same result when monetary policy follows a simple feedback rule in a model expanded by the household heterogeneity of Bilbiie (2024).

Figure 1 also reveals differences generated by the solution techniques. The welfare

---

[16]Naturally, we do this only for the case of the demand shock and occasionally binding lower bound, as the model is otherwise linearized.

[17]A description of both algorithms is provided in the appendix, and a more extended elaboration appears in the online appendix (link here).

loss under the global solution is always higher than it is in the case of the local solution. The difference is increasing as the monetary rule approaches IT ($\xi = 1$). At the upper bound, the welfare loss generated by the global solution is almost 50% higher than the one stemming from the local approximation. Overall, considering that the lower bound is the only non-linearity in an otherwise linearized model, the disparities are not trivial.

Figure 1: Optimal averaging period across the grid of ξ
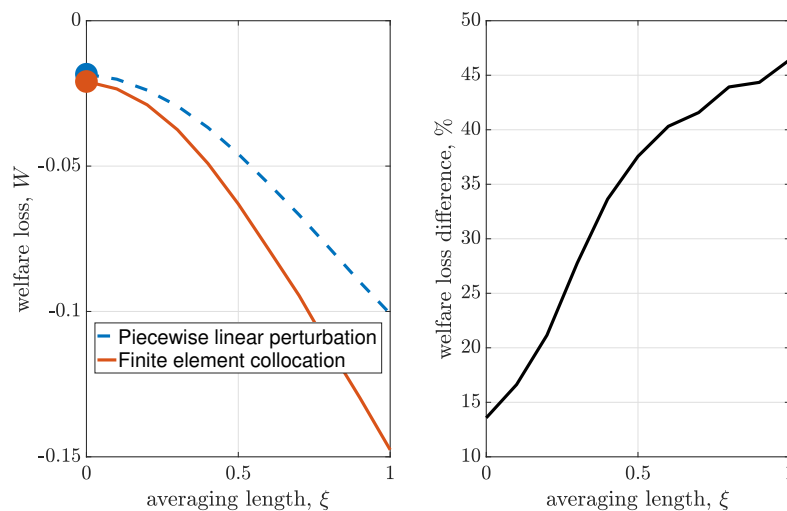
Note: Welfare loss defined by 7 for different parameterizations across a grid of $\xi \in [0,1]$ The generated welfare loss values are multiplied by 100 for ease of interpretation.

Figure 2 shows that a higher degree of makeup behavior generally leads to less time spent at the ELB, regardless of the solution of the model. The time spent at the lower bound is always higher when the global solution is used. Moreover, the difference increases as history-dependence decreases. We stress that the time spent at the lower bound may play a significantly greater role in monetary policy rules analysis than it plays in the model used in this article. Hence, an extension along the lines of modeling some cost mechanism of the zero lower bound could be a valuable advance in future research.

The results highlight the importance of a high degree of rationality for welfare consequences. Even thought the welfare losses decreases monotonically in the window length under both solutions, the true solution using the projection algorithm provides substantially different perception of the costs related to the ELB. On the other

hand, one may question to what degree is the mechanism behind the higher welfare loss in the global solution realistic.[18] The cognitive discounting of Gabaix (2020) enables to explore various degrees of strength of the mechanism without necessarily choosing one of the extreme cases between fully rational agents perfectly internalizing the risk of the ELB in their expectations on one side and perfect foresight agents entirely ignoring it on the other side. To the best of our knowledge, we are the first to analyze such an exogenous risk entering agents' expectations across different degrees of the cognitive discounting of Gabaix (2020).[19]

Figure 2: Fraction of time spent at the ELB across the grid of ξ
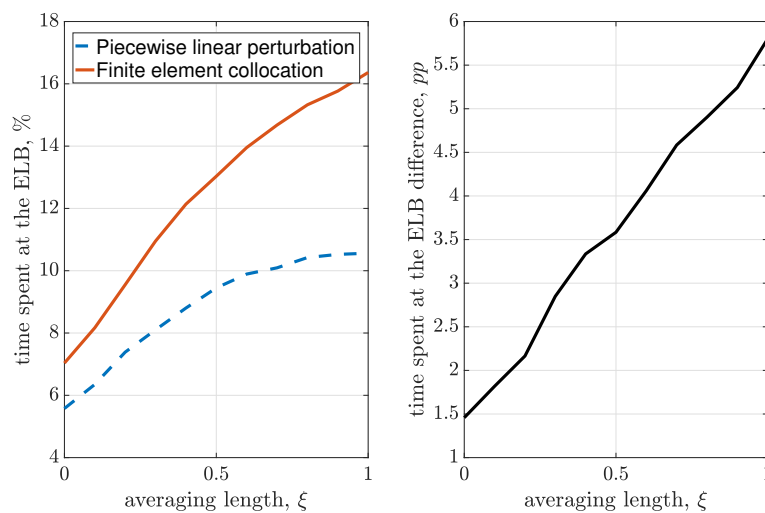
Note: Time spent at the ELB for different parameterizations across a grid of $\xi \in [0, 1]$.

---

[18]Let us repeat that the difference comes from the possibility of hitting the ELB in the future enters the agents expectations which lowers their inflation expectations and results in greater variances of inflation and output gap as the consequence of longer time spent at the lower bound. The mechanism is completely absent under the perfect foresight (which would correspond to $\overline{m} = 0$ with respect to the exogenous risk).

[19]We link these results to Fernández-Villaverde et al. (2023). They show that ELB can lead to higher precautionary savings and hence a lower natural rate of interest if it enters households expectation operator. Even though Fernández-Villaverde et al. (2023) use a HANK model with a non-degenerate distribution and the deflationary bias comes from the natural rate while in our case due to the lower inflation expectations component, they both manifest in a lower policy rate (giving less space before hitting ELB). We point out that also results of Fernández-Villaverde et al. (2023) might non-negligibly change if deviating from the rational expectations equilibrium. We show in our paper substantial quantitative differences of deflationary bias depending on the degree of deviation from rational expectations.

Various calibration of the cognitive discounting. Coibion et al. (2023) show that US households mostly do not understand the change of monetary policy strategy to the AIT rule. Therefore, it is desirable to see how the results may differ under greater myopia. In figure 3, we run the simulations again, but this time we vary the values of the cognitive discounting.[20][21]

Figure 3: Optimal averaging period within the Behavioral THANK model for varying $\overline{m}$
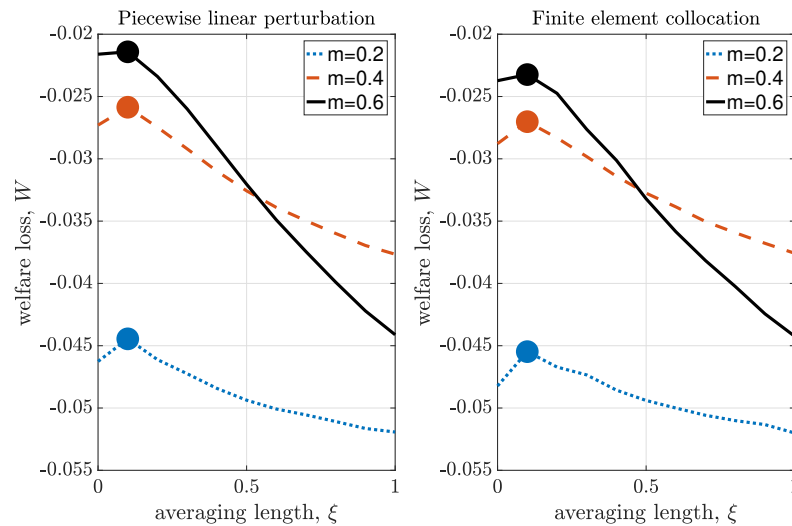
Note: Welfare loss defined by 7 for different values of $\overline{m}$ within the Behavioral THANK model. The generated welfare loss values are multiplied by 100.

We see that even unreasonably small values of the cognitive discounting parameter do not deliver superiority of IT over some form of makeup property. For smaller values of $\overline{m}$, PLT is no longer the best option, yet quite strong history dependence still generates the lowest welfare loss. Even though a difference towards IT is distinctly smaller. Thus, interpretation of our results may go in two directions. One can stress the fact that the superiority of PLT and AIT considerably diminishes in magnitude. Nevertheless, the makeup rules result in lower welfare losses even under strong degree of myopia.

---

[20]For the sake of simplicity, we always keep equivalent parameterization of the general $\overline{m}$ and specifically related m_π^f and m_y^f; i.e., we ignore possible dispersion in the general and concrete discounting parameters.

[21]We keep the variance of the shock the same although lower values of the subjective discounting parameter affect the frequency of hitting the lower bound. Another approach might have been to adjust the strength of the shock in order to keep the time spent at the lower bound constant. However, such a change does not upend the results of the analysis, as shown in Budianto et al. (2023).

Our results differ considerably from those presented in Budianto et al. (2023) for the case of lower values of $\overline{m}$. In their analysis the higher myopia always decreases the welfare loss, while in our case it no longer holds. Figure 3 shows that $\overline{m} = 0.2$ leads to a higher welfare loss than $\overline{m} = 0.4$ across the whole grid of ξ, regardless of the solution method.

More importantly, our analysis shows a higher optimal degree of history dependence under stronger myopia than those of Budianto et al. (2023). In their case, a lower $\overline{m}$ results in a significantly shorter optimal window length, while in our case even very low values of $\overline{m}$ do not move the optimal period away from long-lasting history-dependence. The difference is caused by different monetary policy settings. While Budianto et al. (2023) allow the central bank to use optimal discretion, we let the central bank to follow the feedback rule, in line with Taylor (1993). Our results show that benefits from the history-dependence remain even with stronger subjective discounting when the central bank follows a Taylor-type rule.

This time, the disparities between the solution techniques are substantially less considerable than for the benchmark parameterization. What is more, IT does not deliver a substantially higher discrepancy in the solution methods compared to PLT, as is the case with the benchmark value of $\overline{m}$. The reason behind the convergence of the solution techniques is that the differences were driven by the treatment of expectation operators. Thus, if we allow agents to see the future dimly and partly disregard the risk of the future binding ELB, the two solutions begin to be similar.

3.2 The speed of return to the target path

Our results are not in line with the conclusions of Budianto et al. (2023), who show that moving further from rational expectations leads to less history dependence in the optimal window length. This does not happen in our case. The explanation lies in the behavior of monetary policy and its effect on the speed of return of average inflation to the target path.

Note that Budianto et al. (2023) work with a central bank that conducts monetary policy under optimal discretion. Given that the central bank chooses every period inflation that gets average inflation back to the target, they implicitly assume that the central bank wants to close the gap between average inflation and its target path in-

stantly.

We work with a feedback rule calibrated for a realistic value of the response parameter of the policy rate on the deviations of average inflation from the target. Given that we stick with $\phi_{\pi a} = 1.5$ across the whole parameter space of the window length parameter ξ, we allow the central bank to take some time to return average inflation to the targeted level.[22][23] Put differently, to enforce the same monetary policy with respect to the speed of returning to the target of average inflation as in Budianto et al. (2023), we would need to work with notably higher values of $\phi_{\pi a}$.

We show that our results become qualitatively consistent with Budianto et al. (2023) when the Taylor rule parameter rises to markedly higher values. Specifically, when agents are strongly myopic, and the central bank pushes average inflation back to the target path quickly, IT begins to generate lower welfare loss than PLT, and very short-lasting AIT becomes welfare optimal. For illustration, consider a stronger cognitive discounting of $\overline{m} = 0.4$ and repeat the simulations with different values of $\phi_{\pi a} = \{0.5, 1.5, 5, 10, 50, 100\}$.[24] The welfare analysis for each of the calibrations is shown in figure 4 below.[25]

It is clear that only very (unrealistically) high values of the Taylor rule parameter allow IT to outperform PLT and lead to ξ close to 1 to be optimal. Note that values used in this experiment are sharply inconsistent with the response functions of central banks.[26]

---

[22]Dobrew et al. (2023) optimize the Taylor rule coefficient separately for IT, PLT, and AIT regimes. As noted earlier, we purposefully do not optimize the elasticity parameter in the Taylor rule for each window length, because such a comparison would not consider the same returning time to the target path for different window lengths. What is more, optimizing $\phi_{\pi a}$ would lead to unrealistically high values. Our goal is to keep the monetary policy under AIT close to those studies of central banks in which monetary policy reacts with the same strength regardless of window lengths (see Feiveson et al., 2020; Arias et al., 2020; Hebden et al., 2020).

[23]In figure 10 in the appendix, we highlight how different values of the Taylor rule parameter maps into the distinct speed of return to the target path if the central bank targets average inflation.

[24]We hold $\phi_y = 0.125$ for simplicity as it does not qualitatively upend results.

[25]We use the piecewise linear solution for this purpose. The global approximation would not qualitatively change the results; they would not differ much even quantitatively, given that the risk of hitting the ELB in the future is strongly discounted (see figure 3).

[26]Bear in mind that models inside of central banks usually also work with high values of smoothing parameter in the Taylor rule, while we calibrate it to zero for the sake of economizing on the number of state variables. This would make the response function even more sluggish.

Figure 4: Optimal averaging period for varying $\phi_{\pi a}$
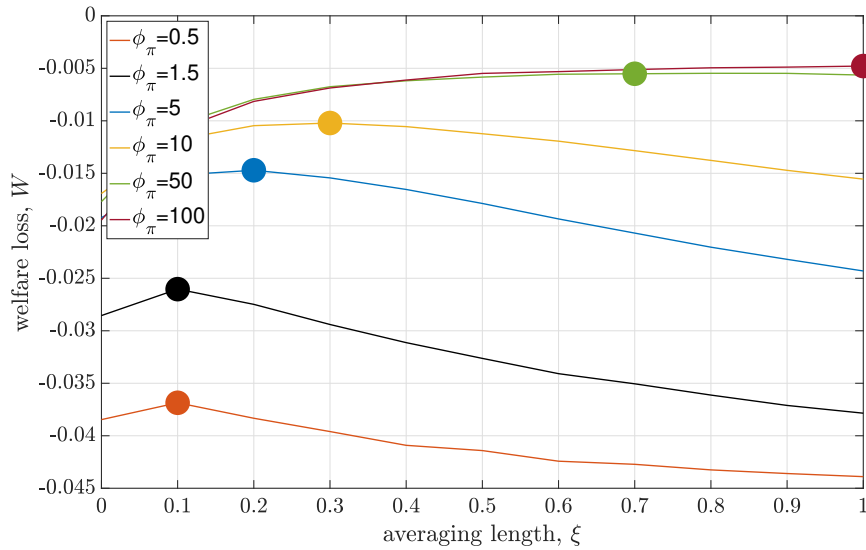
Note: Welfare loss defined by 7 for $\overline{m} = 0.4$ across different values of $\phi_{\pi a}$. The generated welfare loss values are multiplied by 100.

We argue that analysis of the optimal window length in the AIT regime should not be done without explicitly tackling the problem of the speed of return to the target. In other words, how fast the central bank wants to come back to the average inflation target path must be taken into account when evaluating welfare consequences of different averaging periods. As we show, different assumptions about the patience of the central bank with the return to the target path can have decisive effects on the comparison of different window lengths.

The relative benefits of a stronger reaction under IT. We show that IT becomes superior to PLT, while short-lasting AIT is optimal only when the policy rate reacts unrealistically sharply. As figure 4 shows, IT begins to outperform makeup rules, but not due to substantially worse functioning of PLT and AIT. It is rather a lower welfare loss under IT once the central bank starts to react with high sensitivity that drives IT to be superior over long-lasting window lengths.

To highlight this, we show the welfare loss under $\xi = \{0, 1\}$ for $\phi_{\pi a} = \{1.5, 5, 10, 50, 100\}$. We compare the two extreme cases of AIT - PLT and IT - while varying the coefficient in the Taylor rule, and thereby the speed of return to the target. We show how the relative performance of PLT and IT differ depending on the fully rational expectations

formation ($\overline{m} = 1$) and when we impose a strong myopia ($\overline{m} = 0.4$).

Figure 5: Welfare loss for different $\phi_{\pi a}$ and $\overline{m}$ under PLT and IT
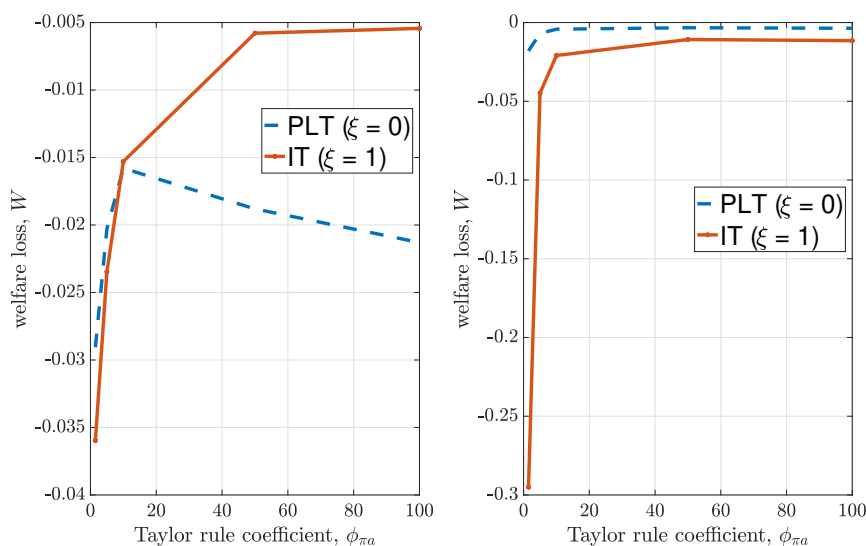
Notes: Welfare implications of the higher speed of return to the target under both IT and PLT for different degrees of myopia ($\overline{m} = 0.4$ on the left and $\overline{m} = 1$ on the right) We use the piecewise linear solution of Guerrieri and Iacoviello (2015).

In the rational expectations case, the PLT rule always generates a lower welfare loss than IT, regardless of the value of the Taylor rule coefficient. However, it is noteworthy that a great deal of difference diminishes when we depart from the values of the coefficient that are commonly used. This is because the brisk reaction function is incorporated into the expectations of agents, which makes the ELB generally much less harmful.[27] Makeup rules can still bring some additional welfare benefits, but these are relatively low compared to the usual Taylor rule calibration.

The comparison starts to become more tangled when we address cognitive discounting. IT again improves its performance more than PLT with increasing $\phi_{\pi a}$. However, because this time the initial difference (when $\phi_{\pi a} = 1.5$) between the two regimes is lower (as PLT is less efficient due to myopic expectations), the higher relative marginal benefits of a stronger reaction in the Taylor rule eventually result in IT being superior to PLT. The better performance of IT under stronger myopia is not a consequence of

---

[27]It is well-known that in a reaction to demand shocks, it is optimal to respond infinitely strongly to deviations of inflation from the target (see Boehm and House, 2014). The presence of the ELB even amplifies this result.

worse operating PLT, but rather large benefits of sharper response of the Taylor rule under IT.

The source of IT superiority. It may be tempting to interpret the results of Budianto et al. (2023) as indicating that stronger cognitive discounting erodes the stabilization expectations channel embedded in makeup rules. The mechanism relying on the rational expectations formation goes back to Svensson (1999) and has since been highlighted in the literature concerning ELB. Naturally, if we deviate from rational expectations by invoking stronger cognitive discounting, we could assume that the inability to shape expectations sufficiently would lead to sharper policy rate responses, resulting in greater variance of output and inflation and ultimately higher welfare loss. The counterargument against makeup rules that was pervasive before Svensson (1999) would appear to be valid. Our results challenge such reasoning, because we show long-lasting averaging to also be optimal under stronger myopia.

While 4 reveals that optimal averaging comes close to IT only if we work with high values in the Taylor rule, 5 clarifies that it is not a deterioration of welfare loss under long-lasting averaging, but rather the benefit of stronger Taylor rule elasticity that delivers shorter windows to be optimal under more severe cognitive discounting.

By disentangling the difference between the results of Budianto et al. (2023) and ours, we highlight that it is necessary to take the speed of return to the target path into account, as figure 5 shows that, at some point, a faster return (higher Taylor rule parameter) in the PLT rule starts to generate greater welfare loss. This holds for both rational expectations ($\overline{m} = 1$) and a substantial myopia ($\overline{m} = 0.4$) cases (although it is not well-distinguishable on the right part of 5 due to the scale), even though it is more pronounced in the latter situation.

4 Extensions of the baseline model

4.1 A backward-looking component in the expectations

We extend our analysis to the case of expectation formation with backward-looking components. Although we work with a boundedly rational approach to attenuate the strength of the expectation channels, we assume that agents do not pay any attention

to past outcomes in the economy. However, such an assumption may be seen as too strong. For instance, Malmendier and Nagel (2015) show that past inflation plays a substantial role in forming expectations. Using an adaptive learning model, they estimate the weight of previous inflation in the expectation formation process to be 0.6.[28] Given that we study a make-up monetary policy rule, we extend the model with a backward-looking component only for inflation.

Assume that, this time, households do not shrink their forecast toward the steady state, but rather toward the past inflation outcome. Hence, instead of $\mathbb{E}_t^{BR}[\pi_{t+1}] = \overline{m}\,\mathbb{E}_t[\pi_{t+1}]$, we now work with $\mathbb{E}_t^{BR}[\pi_t] = \overline{m}\,\mathbb{E}_t[\pi_{t+1}] + (1 - \overline{m})\pi_{t-1}$. Consequentially, we end up with a different dynamic IS curve of the following form:

$$\tilde{y}_t = \psi_f\,\mathbb{E}_t\,\tilde{y}_{t+1} - \psi_c\frac{1}{\sigma}\left\{i_t - \left[\overline{m}\,\mathbb{E}_t\,\pi_{t+1} + (1 - \overline{m})\pi_{t-1})\right] - r_t^{nat}\right\} \tag{9}$$

Moving to the supply side, we follow Christiano et al. (2005) and assume that firms that are unable to reset their prices apply full indexation to the previous inflation rate. Following Gabaix (2020), we can write the NK PC in the deviation of inflation from its previous value (hence assuming the past inflation to be the default value):

$$\hat{\pi}_t = \beta M^f\mathbb{E}_t\hat{\pi}_{t+1} + \kappa\tilde{y}_t, \tag{10}$$

where $\hat{\pi}_t = \pi_t - \pi_{t-1}$. Substituting into equation 10 and reshuffling yields:

$$\pi_t = (1 + \beta M^f)^{-1}(\beta M^f\mathbb{E}_t\pi_{t+1} + \pi_{t-1} + \kappa\tilde{y}_t) \tag{11}$$

Note that equation 11 is closely related to the canonical backward-looking PC approach of Galí and Gertler (1999). An isomorphic equation can be also derived using the incomplete information perspective of Angeletos and Lian (2018).

We run the same simulations as we do in section 3 when we work with the benchmark model. However, this time we use equations 9 and 11 to incorporate partially backward-looking inflation expectations, and compare the behavior of this model with the original.[29] The results for the baseline calibration of the cognitive discounting are depicted in figure 6 below.[30]

---

[28]Consider a convex combination leading to resulting inflation expectations.

[29]We use only the global solution for this purpose, because the general meaning of the comparison of solution techniques would not be altered.

[30]We do not show simulations for lower values of the subjective discounting parameter in the main

Figure 6: Optimal averaging period - comparison with a backward-looking model
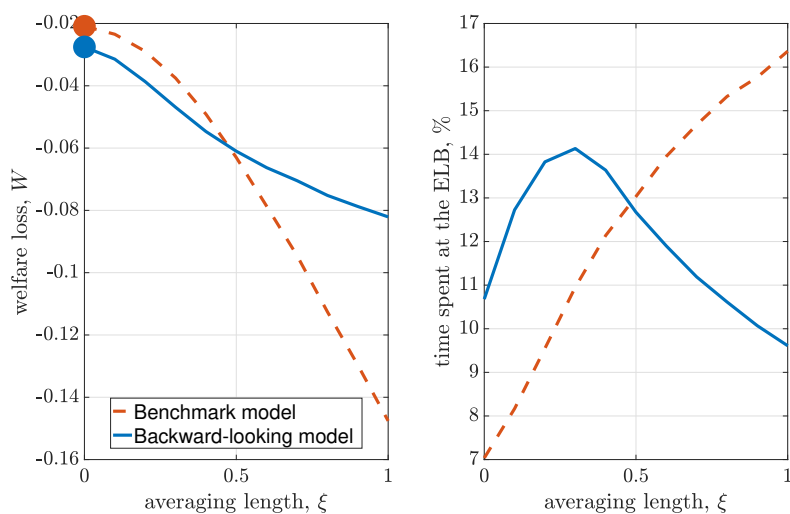
Note: Welfare loss defined by 7 and time spent at the ZLB for different parameterizations across a grid of $\xi \in [0, 1]$. The red curves depicts the benchmark model and the blue curves are the model extended by the backward-looking component for inflation expectations. The generated welfare loss values are multiplied by 100 for ease of interpretation.

The figure shows some important distinctions. The backward-looking component on the inflation expectations moderates benefits stemming from the history-dependent characteristic in the monetary policy rule. The difference between extreme situations - IT and PLT - is markedly lower when agents take into account the previous outcomes of inflation. Hence, the presence of backward-looking agents can help to alleviate the burden of the ZLB, because agents do not project deteriorating economic conditions due to understating of inability to further adjust real interest rate when the central bank targets inflation. Simultaneously, we observe that the economy spends significantly less time at the zero lower bound under IT than in the benchmark model. The opposite is true when the monetary rule embeds long-lasting averaging. The central bank needs to keep the nominal rates at zero longer, given that the inflation expectation channel that further lowers the real interest rate in the fully forward-looking model is weakened here. Moreover, the welfare loss under PLT is greater with the backward-

text, because they do not differ considerably from the benchmark case. This is because a lower value of $\overline{m}$ itself attenuates the expectation channel via which make-up rules should benefit the ZLB situation. Therefore, including backward-looking inflation expectations does not deliver stark differences as comparing the baseline case of $\overline{m} = 0.85$.

looking component. The reason is again the fact that the expectation channel through which make-up rules may help by changing inflation expectations based on deviations from the target path relies on a forward-looking structure.

Nevertheless, despite the reasons noted for reducing the virtues of the history-dependent feature, the general meaning of the results do not change. The infinite averaging still generates the lowest welfare loss, and the loss is monotonically decreasing in the averaging length. These main outcomes are present in both the benchmark and in the partially backward-looking models.

4.2 Quantitative easing

Up to now, we have assumed that the central bank has only one monetary policy tool available to handle shocks buffeting the economy. However, relying on the assumption that monetary policy has no additional instruments besides the short-term nominal interest rate may skew analysis of makeup monetary policy rules in their favor relative to IT. The ELB situation may still be too painful when the central bank cannot implement unconventional monetary policy tools when the economy hits the lower bound and the central bank targets inflation. Naturally, the difference between the welfare loss under PLT or AIT and IT may be overly high as the lower bound leads to a substantial loss in IT.

Severity of the lower bound stemming from the binding lower bound within a New Keynesian framework under the IT rule proved to be at odds with the real situation after the Great Recession. Debortoli et al. (2019) show that the implications of the severity of the ELB are inconsistent, when they compare the conventional New Keynesian model with the reality observed after the Great Recession, when many economies remained at the lower bound for prolonged time periods.

In addition to the unrealistic costs of the binding lower bound, New Keynesian models suffer from some peculiar paradoxes that appear in the ELB situation. Eggertsson (2010); Eggertsson and Krugman (2012); Bhattarai et al. (2018); Christiano et al. (2011) show how the presence of ELB in the model delivers features including the paradox of flexibility, the paradox of toil, or unreasonably large values of the fiscal multiplier. As a response to these issues, Bonciani and Oh (2021) point out that extending a simple canonical three equation New Keynesian model by quantitative easing solves

all of the paradoxes at once. Particularly, Bonciani and Oh (2021) use the four equation version of the simple New Keynesian model of Sims et al. (2021), which does not suffer from these problems thanks to its extension by the unconventional monetary policy.

We use the model of Sims et al. (2021) and merge it with the Behavioral THANK model of Pfäuti and Seyrich (2023), including the quantitative easing (QE) policy into Pfäuti and Seyrich (2023). The altered dynamic IS curve looks as follows:[31]

$$\tilde{y}_t = \psi_f \, \mathbb{E}_t \, \tilde{y}_{t+1} - (1-z)\psi_c \frac{1}{\sigma}(i_t - \overline{m} \, \mathbb{E}_t \, \pi_{t+1} - r_t^{nat}) - z \frac{b^{CB}}{b}(\psi_f \, \mathbb{E}_t \, qe_{t+1} - qe_t), \quad (12)$$

where the notation related to quantitative easing follows Sims et al. (2021). Thus, z is the share of impatient (child) agents who always borrow on their consumption, $\frac{b^{CB}}{b}$ denotes the fraction of long-term bonds held by the central bank and, most importantly, qe_t is the real market value of the bond portfolio that the central bank holds.[32] Note that assigning $z = 0$ boils the model back down to 1.

In addition to its effect on the demand side, unconventional monetary policy also affects the supply sector through marginal costs. Hence, the PC has the following form:

$$\pi_t = \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa^{QE} \tilde{y}_t - m_y^f \omega \sigma \frac{z}{(1-z)} \frac{b^{CB}}{b} qe_t, \quad (13)$$

where $\kappa^{QE} = m_y^f \overline{\kappa}^{QE}$ which is isomorphic to the case without the QE. However, this time $\overline{\kappa}^{QE} = \omega \left[\frac{\varphi(1-z)+\sigma}{(1-z)} \right]$. All the new parameters have the same meaning as in the IS curve. Equation 2 is again a special case of 13 when $z = 0$.

Clearly, the model of Sims et al. (2021), which can be retrieved from 12 and 13 by setting $\lambda = 0$ and $\overline{m} = m_y^f = m_\pi^f = 1$, is unable to simultaneously resolve both issues that Pfäuti and Seyrich (2023) handle. The four equation model of Sims et al. (2021) actually worsens the amplification of the contemporary monetary policy shock, while leaving untouched the problem of the forward guidance puzzle. Moreover, when we use the baseline parameterization from the previous simulations and mix it with the calibration used in Sims et al. (2021) ($z = 1/3$ and $\frac{b^{CB}}{b} = 0.3$), we lose the amplification

---

[31]The derivation steps of merging Sims et al. (2021) and Pfäuti and Seyrich (2023) are in the appendix and in detailed form in the online appendix (link here). We ignore the credit shocks that are present in Sims et al. (2021).

[32]Note that now we have both the heterogeneity coming from Bilbiie (2024) and the heterogeneity related to Sims et al. (2021). We merge these two into our final equations presented in the main text in the appendix. The online appendix provides a full derivation of the QE extended model (link here).

feature even in our equation ($(1-z)\psi_c\frac{1}{\sigma} = 0.8739$).[33] The forward guidance puzzle resolution still holds. The sensitivity to quantitative easing is $z\frac{b^{CB}}{b} = 0.1$.

To maintain comparability of the QE extension to the initial model, we run the welfare analysis with two versions of the extended model. In one of them, we follow the parameterization of Sims et al. (2021) and work with a fully micro-founded model. In the second case, we keep the original model with ψ_c and κ only adding ad hoc the quantitative easing components from equations 12 and 13 with the parameterization from Sims et al. (2021).[34]

The central bank now uses the purchase of long-term bonds as a monetary policy tool in addition to the short term policy rate.[35] Specifically, we follow Bonciani and Oh (2021) and set the rule for the QE as:[36]

$$qe_t = \begin{cases} -\upsilon\pi_t & \text{if } i = 0 \\ 0 & \text{if } i = \rho_i i_{t-1} + (1-\rho_i)\left[(\frac{1}{\beta}-1) + \phi_{\pi a}\pi_t^a + \phi_y \tilde{y}_t\right] > 0 \end{cases} \tag{14}$$

where υ drives the sensitivity of the bond purchase on the inflation deviations from the target. Bonciani and Oh (2021) show that, to get rid of the paradox of flexibility under the central bank following the Taylor rule, $\upsilon \geq 28$ is necessary.[37] Consequentially, we set a similar value $\upsilon = 30$ in our analysis. As is clear from the piecewise equation 14, the QE is activated only at the policy-rate lower bound.

---

[33]To preserve the amplification characteristic, we need to have sufficiently high χ. Specifically, the following must hold: $\chi > \frac{1-(1-z)(1-\lambda)}{\lambda}$. In the case of our baseline parameterization, we need to have $\chi > 1.6746$. Thus, if we decrease the share of children among the households, we still have the amplification, even though in a weaker form. Specifically, $z < 1 - \frac{(1-\lambda\chi)}{(1-\lambda)}$ is a necessary to hold for the amplification. In our baseline parameterization, this means $z < 0.2372$. Note that this calibration of z also results in a weaker effect of the QE. One can use a higher value of $\frac{b^{CB}}{b}$ as a push-back to keep the strength of the unconventional monetary policy unchanged.

[34]Hence, we have $\tilde{y}_t = \psi_f \mathbb{E}_t \tilde{y}_{t+1} - \psi_c \frac{1}{\sigma}(i_t - \overline{m}\mathbb{E}_t \pi_{t+1} - r_t^{nat}) - z\frac{b^{CB}}{b}(\psi_f \mathbb{E}_t qe_{t+1} - qe_t)$ and $\pi_t = \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa\tilde{y}_t - \omega\sigma\frac{z}{(1-z)}\frac{b^{CB}}{b}qe_t$. This ad-hoc system is used only for a direct comparison to Pfäuti and Seyrich (2023) without QE.

[35]As Sims et al. (2021) show, equation 12 can be reshuffled to show the term spread of the yield curve by substituting for qe. Nevertheless, this is at the expense of less tractability of the model. We could not write the model as the four-equation system anymore because we would need to keep track of more variables.

[36]Bonciani and Oh (2021) point out the importance of connecting the QE policy with inflation to rule out the NK paradoxes at the ELB.

[37]The other two paradoxes are related to fiscal policy which we do not address in our analysis.

We run the welfare analysis under the natural rate of interest shock and the ELB presence in the same way as we did before.[38] However, this time we allow for the central bank to deploy quantitative easing every time the economy hits the lower bound while the QE reacts to inflation.

Figure 7: Optimal averaging period across the grid of ξ taking into account the QE
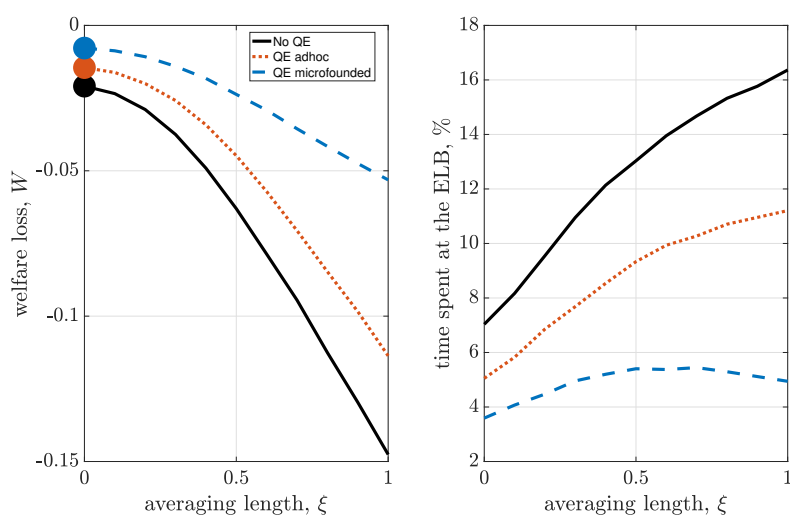
Note: Welfare loss for the benchmark parameterization for the natural rate of interest shock when taking the QE into account. The generated welfare loss values are multiplied by 100.

Figure 7 shows that the results depend on how we treat parameterization of the QE-extended model of Pfäuti and Seyrich (2023). When we use the ad-hoc approach, the presence of the unconventional monetary policy does not change much compared to the model without the QE. We can see that the presence of the unconventional monetary policy decreases the welfare losses across the whole grid of ξ, but it does not shape the relative ordering of different history-dependence strengths in any manner.

When we use parameterization consistent with the micro-founded equations 12 and 13, the welfare losses are substantially attenuated. This is because the QE extension of Sims et al. (2021) calibrated in line with the original article and mixed with our baseline parameterization of Pfäuti and Seyrich (2023) no longer delivers the amplification feature. This also means that the ELB is less painful, because the amplification also implies a more costly lower bound compared to the benchmark RANK case.

---

[38]Note that we deploy only the true solution delivered by the global approximation in this case. The story of the comparison of the solution techniques would not differ from the original model without the QE. We felt it would be redundant to repeat it in this extension.

Interestingly, figure 7 shows that, when the central bank can use the QE, the combination of unconventional monetary policy and history-dependence results in a remarkable drop in the time spent at the lower bound. In total, inclusion of the unconventional monetary policy may deliver some cosmetic changes relative to the simple monetary policy case, yet it does not fundamentally upend anything. Most importantly, the optimal window length under $\overline{m} = 0.85$ is still infinite. We also ran equivalent simulations for lower values of the myopia parameter, and the core results do not vary substantially with respect to the baseline case. Hence, we do not include them in the text for the sake of conciseness.

4.3 Supply shocks

We have analyzed the optimal averaging window length considering only the demand shock and the ELB presence. We deem to analyze supply shocks of less interest for various reasons. Firstly, the change that the Federal Reserve made in August 2020 was based on the ELB argument, as documented in Clarida (2020) and Clarida (2023). Given that it was never the central bank's intention to undo previous deviations from the target path caused by supply fluctuations, nor to correct previous deviations above the target path in any way, studying supply driven fluctuations in relation with AIT seems less important.[39]

The results of such exercises are easy to anticipate. Gabaix (2020) shows that price level targeting dominance in response to mark-up shocks breaks down when firms are boundedly rational. Hence, it is straightforward to expect that the optimal window length shortens with increasing firms myopia. Budianto et al. (2023) show exactly such results in their article.

For the sake of completeness, we also conduct an analysis with supply shocks, but we keep the discussion brief in the main text and include more detail in the online appendix. In addition to incorporating supply shocks buffeting the economy as mark-up shocks defined by the AR(1) process, we also adjust the supply shock by persistence on the innovations, and therefore use an ARIMA(1,1,3) process following Walsh (2022). The purpose of such a defined shock process is to investigate the AIT in the presence

---

[39]In some aspects, the new regime is close to the temporary price level targeting mentioned in Bernanke (2017*b*) and Bernanke (2017*a*). This fact was explicitly acknowledged in Clarida (2021).

of long-lasting persistent problems on the supply side, resembling the years after the Fed's regime change. In the case of the ARIMA process, the welfare loss is considerably higher across the whole grid of ξ, but it does not change the qualitative results. We depict the results without much elaboration in figure 8 above. Our results are in line with Budianto et al. (2023) and the reasoning of Gabaix (2020) given that including heterogeneous agents on the households side, unlike in their articles, does not alter the conclusions.

Figure 8: Optimal averaging period across the grid of ξ for supply shocks
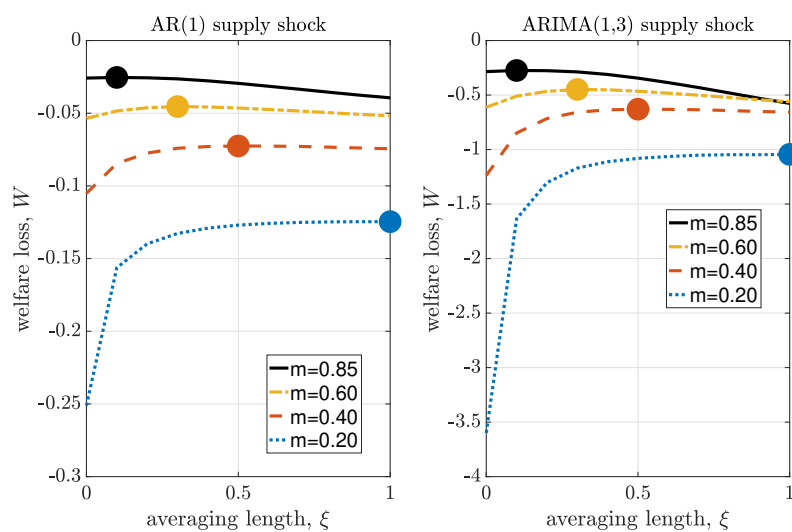
Note: Welfare loss for different values of $\overline{m}$ considering both supply shock types. The generated welfare loss values are multiplied by 100.

5   Conclusion

We consider the problem of the optimal window length in the average inflation targeting regime. We compute the optimal averaging period using a model that can handle two crucial features of monetary policy transmission that a fully rational New Keynesian model cannot deliver. Our model rules out the forward guidance puzzle and can simultaneously generate an amplification of the contemporary monetary policy shock through an indirect general equilibrium effect. Moreover, we allow for the central bank to use unconventional monetary policy in the form of QE. We expand our analysis further by incorporating a backward-looking component of inflation expectations.

We show that the optimal averaging window is infinite for a conventional value of the cognitive discounting parameter from Gabaix (2020); price level targeting dominates both inflation targeting and average inflation targeting. This holds even for extensions of the benchmark model by adding QE and past inflation outcomes. Once the degree of myopia starts rising, the infinite averaging window is dominated by average inflation targeting with a finite history dependence. However, the optimal window length is still long-lasting and does not converge to IT with increasing myopia. We explain this as a result of the strength of the central bank's response to push average inflation back to the target. When the central bank is patient and does not need to push average inflation to the target instantly, PLT and AIT still outperform IT. With a stronger response to return average inflation to the target path quickly, shorter window lengths dominate, and IT eventually becomes superior. Hence, we argue that the question of speed of return must be taken into consideration when evaluating different averaging periods in AIT.

We also disentangle how the mere possibility of a binding lower bound in the future may alter the results. We obtain the magnitude of the uncertainty channel by solving the model twice for the smooth process of exponential MA that economizes on the number of state variables. The piecewise first-order perturbation solution method of Guerrieri and Iacoviello (2015) completely ignores the expectation channel of the future possibility of hitting the lower bound. On the contrary, the finite element approach using collocation and linear splines as basis functions can incorporate the uncertainty and hence the potential precautionary behavior in the model. Observing the difference between these two solution techniques enables us to identify the welfare loss differences generated by this channel. We quantify the magnitude of the deflationary bias (see Eggertsson, 2006; Penalver and Siena, 2024). We show that the differences are non-negligible and substantially increase with low or zero degrees (inflation targeting) of history dependence. The global solution generates a greater welfare loss regardless of the monetary rule. Nevertheless, under price level targeting, the difference shrinks, though it is considerable in inflation targeting. Furthermore, the difference starts to disappear under a stronger cognitive discounting as the uncertainty is discounted more.

Bibliography

Adam, K. and Billi, R. M. (2006), 'Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates', *Journal of Money, Credit and Banking* **38(7)**, 1877–1905.

Adam, K. and Billi, R. M. (2007), 'Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates', *Journal of Monetary Economics* **54(3)**, 728–752.

Amano, R., Gnocchi, S., Leduc, S. and Wagner, J. (2020), 'Average is Good Enough: Average-inflation Targeting and the ELB', *Bank of Canada: Staff Working Paper* .

Angeletos, G.-M. and Lian, C. (2018), 'Forward Guidance without Common Knowledge.', *American Economic Review* **108 (9)**, 2477–2512.

Arias, J., Bodenstein, M., Chung, H., Drautzburg, T. and Raffo, A. (2020), 'Alternative Strategies: How Do They Work? How Might They Help?', *Finance and Economics Discussion Series 2020-068.* .

Auclert, A., Rognlie, M. and Straub, L. (2020), 'Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model', *Working Paper, National Bureau of Economic Research* .

Benigno, G., Hofmann, B., Nuño, G. and Sandri, D. (2024), 'Quo vadis, r*? The natural rate of interest after the pandemic', *BIS Quarterly Review* .

Beqiraj, E., Di Bartolomeo, G. and Di Pietro, M. (2019), 'Beliefs Formation and the Puzzle of Forward Guidance Power', *Journal of Macroeconomics* **60**, 20–32.

Bernanke, B. (2017*a*), 'Monetary Policy in a Low Interest Rate World: Comment', *n Brookings Papers on Economic Activivity. Brookings Institution.* .

Bernanke, B. (2017*b*), 'Monetary Policy in a New Era', *Rethinking Macroeconomic Policy, Peterson Institute* .

Bernanke, B. S. (2000), 'Japanese Monetary Policy: A Case of Self-Induced Paralysis', *Princeton university* .

Bhattarai, S., Eggertsson, G. B. and Schoenle, R. (2018), 'Is Increased Price Flexibility Stabilizing? Redux', *Journal of Monetary Economics* **100**, 66–82.

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

Bilbiie, F. (2024), 'Monetary Policy and Heterogeneity: An Analytical Framework', *Review of Economic Studies* **Forthcoming**.

Billi, R. M. (2017), 'A Note On Nominal GDP Targeting and the Zero Lower Bound', *Macroeconomic Dynamics* **21(8)**, 2138–2157.

Boehm, C. E. and House, C. L. (2014), 'Optimal Taylor Rules in New Keynesian Models', *NBER Working Paper No. 20237* .

Bonciani, D. and Oh, J. (2021), 'Revisiting the New Keynesian Policy Paradoxes under QE', *Bank of England working papers* .

Branch, W. and McGough, B. (2009), 'A New Keynesian Model with Heterogeneous Expectations', *Journal of Economic Dynamics and Control* **33**(5), 1036–1051.

Budianto, F., Nakata, T. and Schmidt, S. (2023), 'Average Inflation Targeting and the Interest Rate Lower Bound', *European Economic Review* **152**, 104384.

Bäcker-Peral, V., Hazell, J. and Mian, A. (2024), 'Dynamics of the Long Term Housing Yield: Evidence from Natural Experiments', *mimeo* .

Calvo, G. (1983), 'Staggered Prices in a Utility-Maximizing Framework', *Journal of Monetary Economics* **12**(3), 383–398.

Christiano, L., Eichenbaum, M. and Evans, C. L. (2005), 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy', *Journal of Political Economy* **113(1)**, 1–45.

Christiano, L., Eichenbaum, M. and Rebelo, S. (2011), 'When Is the Government Spending Multiplier Large?', *Journal of Political Economy* **119(1)**, 78–121.

Clarida, R. H. (2020), 'The Federal Reserve's New Monetary Policy Framework: A Robust Evolution', *Speech 88667, Board of Governors of the Federal Reserve System (U.S.)* .

Clarida, R. H. (2021), 'The Federal Reserve's New Framework and Outcome-Based Forward Guidance', *Speech at "SOMC: The Federal Reserve's New Policy Framework" a forum sponsored by the Manhattan Institute's Shadow Open Market Committee, New York.* .

Clarida, R. H. (2023), 'US Monetary Policy and the Return to Price Stability', *National Bureau of Economic Research* .

Coibion, O., Gorodnichenko, Y., Knotek II, E. S. and Schoenle, R. (2023), 'Average Inflation Targeting and Household Expectations', *Journal of Political Economy Macroeconomics* **1:2**, 403–446.

Coibion, O., Gorodnichenko, Y., Kueng, L. and Silvia, J. (2017), 'Monetary Policy, Income Inequality, Consumption Inequality', *Journal of Monetary Economics* **88**, 70–89.

Coulter, J., Duncan, R. and Martínez-García, E. (2022), 'Flexible Average Inflation Targeting: How Much Is U.S. Monetary Policy Changing?', *Globalization Institute Working Paper, Federal Reserve Bank of Dallas* .

Debortoli, D., Galí, J. and Gambetti, L. (2019), 'On the Empirical (Ir)Relevance of the Zero Lower Bound Constraint', *in Martin S. Eichenbaum, Erik Hurst, and Jonathan A. Parker, eds., NBER Macroeconomics Annual, The University of Chicago Press* **34**, 141–170.

Del Negro, M., Giannoni, M. and Patterson, C. (2013), 'The Forward Guidance Puzzle', *Federal Reserve Bank of New York Staff Report 574* .

Djeutem, E., He, M., Reza, A. and Zhang, Y. (2022), 'Household Heterogeneity and the Performance of Monetary Policy Frameworks', *Bank of Canada: Staff Working Paper* .

Dobrew, M., Gerke, R., Kienzler, D. and Schwemmer, A. (2023), 'Monetary Policy Rules under Bounded Rationality', *Deutsche Bundesbank Discussion Paper No. 18/2023* .

Eggertson, G. B. and Woodford, M. (2003), 'The Zero Bound on Interest Rates and Optimal Monetary Policy', *Brookings Papers on Economic Activity 2* **34**(1), 139–211.

Eggertsson, G. B. (2006), 'The Deflation Bias and Committing to Being Irresponsible', *Journal of Money, Credit and Banking* **38 (2)**, 283–321.

Eggertsson, G. B. (2010), 'The Paradox of Toil', *FRB of New York Staff Report No. 433.* .

Eggertsson, G. B. and Krugman, P. (2012), 'Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach', *Quarterly Journal of Economics* **127(3)**, 1469–1513.

Fed (2020), 'Review of Monetary Policy Strategy, Tools, and Communications', https://www.federalreserve.gov/monetarypolicy/review-of-monetary-policy-strategy-tools-and-communications.htm.

Feiveson, L. J., Gornemann, N., Hotchkiss, J. L., Mertens, K. and Sim, J. W. (2020), 'Distributional Considerations for Monetary Policy Strategy', *Finance and Economics Discussion Series 2020-073, Board of Governors of the Federal Reserve System (U.S.)* .

Fernández-Villaverde, J., Gordon, G., Guerrón-Quintana, P. A. and Rubio-Ramírez, J. (2015), 'Nonlinear Adventures at the Zero Lower Bound', *Journal of Economic Dynamics and Control* **57**, 182–204.

Fernández-Villaverde, J., Marbet, J., Nuño, G. and Rachedi, O. (2023), 'Inequality and the zero lower bound', *National Bureau of Economic Research, Working Paper Series 31282* .

Gabaix, X. (2020), 'A Behavioral New Keynesian Model', *American Economic Review* **110**(8), 2271–2327.

Galí, J. (2015), *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition.*

Galí, J. and Gertler, M. (1999), 'Inflation Dynamics: A structural Econometric Analysis', *Journal of Monetary Economics* **44**(2), 195–222.

Guerrieri, L. and Iacoviello, M. (2015), 'Occbin: A Toolkit For Solving Dynamic Models With Occasionally Binding Constraints Easily', *Journal of Monetary Economics* **70 (C)**, 22–38.

Hebden, J., Herbst, E. P., Tang, J., Topa, G. and Winkler, F. (2020), 'How Robust Are Makeup Strategies to Key Alternative Assumptions?', *Finance and Economics Discussion Series 2020-069.* .

Honkapohja, S. and McClung, N. (2021), 'On Robustness of Average Inflation Targeting', *Research Discussion Papers 6/2021, Bank of Finland* .

Jia, C. and Wu, J. C. (2021), 'Average Inflation Targeting: Time Inconsistency and Intentional Ambiguity', *Working Papers 21-19R, Federal Reserve Bank of Cleveland, revised 01 Feb 2022* .

Kaplan, G., Moll, B. and Violante, G. L. (2018), 'Monetary Policy According to HANK', *American Economic Review* **108**(3), 697–743.

Krugman, P. R. (1998), 'It's Baaack: Japan's Slump and the Return of the Liquidity Trap', *Brookings Papers on Economic Activity* **29**(2), 137–205.

Malmendier, U. and Nagel, S. (2015), 'Learning from Inflation Experiences', *Quarterly Journal of Economics* **131 (1)**, 53–87.

Martínez-García, E. (2021), 'Get the Lowdown: The International Side of the Fall in the U.S. Natural Rate of Interest', *Economic Modelling* **100**, 105486.

McKay, A., Nakamura, E. and Steinsson, J. (2016), 'The Power of Forward Guidance Revisited', *American Economic Review* **106**(10), 3133–3158.

Mertens, T. M. and Williams, J. C. (2019), 'Monetary Policy Frameworks and the Effective Lower Bound on Interest Rates', *AEA Papers and Proceedings* **109**, 427–432.

Miranda, M. J. and Fackler, P. L. (2002), *Applied Computational Economics and Finance*.

Mumtaz, H. and Theophilopoulou, A. (2017), 'The Impact of Monetary Policy on Inequality in the UK. An Empirical Analysis', *European Economic Review* **98**, 410–423.

Nahmias, S. and Olsen, T. L. (2015), *Production and Operations Analysis*.

Nakata, T. and Schmidt, S. (2019), 'Conservatism and Liquidity Traps', *Journal of Monetary Economics* **104**, 37–47.

Nakov, A. (2008), 'Optimal and Simple Monetary Policy Rules with Zero Floor on the Nominal Interest Rate', *International Journal of Central Banking* **4(2)**, 73–127.

Nessén, M. and Vestin, D. (2005), 'Average Inflation Targeting', *Journal of Money, Credit and Banking* **37**(5), 837–863.

Obstfeld, M. (2023), 'Natural and Neutral Real Interest Rates: Past and Future', *National Bureau of Economic Research, Working Paper Series 31949* .

Penalver, A. and Siena, D. (2024), 'The Deflationary Bias of the ZLB and the FED's Strategic Response', *Journal of Money, Credit and Banking* **Forthcoming**.

Pfäuti, O. and Seyrich, F. (2023), 'A Behavioral Heterogeneous Agent New Keynesian Model'.

Piergallini, A. (2022), 'Average Inflation Targeting and Macroeconomic Stability', *Economics Letters* **219**.

Samarlna, A. and Nguyen, A. D. M. (2019), 'Does Monetary Policy Affect Income Inequality in the Euro Area?'.

Schnabel, I. (2024), 'R(ising) star?', *Speech by Isabel Schnabel, Member of the Executive Board of the ECB, at The ECB and its Watchers XXIV Conference session on: Geopolitics and Structural Change: Implications for Real Activity, Inflation and Monetary Policy, mimeo* .

Sims, E. R., Wu, C. J. and Zhang, J. (2021), 'The Four Equation New Keynesian Model', *The Review of Economics and Statistics, Forthcoming* p. 1–45.

Svensson, L. E. O. (1999), 'Price-Level Targeting versus Inflation Targeting: A Free Lunch?', *Journal of Money, Credit and Banking* **31**(3), 277–295.

Svensson, L. E. O. (2001), 'The Zero Bound in an Open Economy: A Foolproof Way of Escaping from a Liquidity Trap', *Monetary and Economic Studies, Institute for Monetary and Economic Studies, Bank of Japan* **19(S1)**, 277–312.

Svensson, L. E. O. (2020), 'Monetary Policy Strategies for the Federal Reserve', *International Journal of Central Banking* **16**, 133–193.

Taylor, J. B. (1993), 'Discretion versus Policy Rules in Practice', *Carnegie-Rochester Conference Series on Public Policy* **39**(1), 195–214.

Walsh, C. E. (2022), 'Inflation Surges and Monetary Policy', *IMES Discussion Paper Series 22-E-12, Institute for Monetary and Economic Studies, Bank of Japan* .

Appendix

A.1   Suitability of the model for makeup rules analysis

The final private sector equations 1 and 2 consist of features that make the model suitable to investigate the makeup monetary policy rules. The superiority of these regimes over IT can often arise from a fully RANK environment. Specifically, an important caveat of working with PLT, AIT, or nGDPLT rules within a New Keynesian framework is rooted in an exaggerated sensitivity of the current output and inflation on the expected path of the real interest rate - the forward guidance puzzle (Del Negro et al., 2013; McKay et al., 2016). Considering that our model can resolve the puzzle thanks to the cognitive discounting of Gabaix (2020) makes it more capable of assessing the makeup regimes within it.

What is more, Kaplan et al. (2018), Bilbiie (2024), and Auclert et al. (2020) show the importance of an indirect general equilibrium effect in the transmission of monetary policy. However, conventional New Keynesian models almost completely abstain from this effect as monetary policy works dominantly through a direct effect that is driven by intertemporal substitution in consumption. A tractable version of the HANK model of Bilbiie (2024) offers an environment in which analysis of makeup regimes is based on more realistic assumptions about the monetary policy transmission channels.

Nevertheless, as Pfäuti and Seyrich (2023) show, one needs to mix both the THANK and Behavioral NK models to lessen too strong reaction of current variables on the future, yet at the same time to deliver monetary policy amplification caused by the indirect channel. Ignoring the myopic features of Gabaix (2020) would pose a trade-off between the resolution of the forward guidance and the monetary policy amplification, resulting in the *Catch-22* (Bilbiie, 2024). The model would generate the former in the case of procyclical consumption inequality, but the stronger reaction to the contemporaneous real interest rate would be missing. Conversely, the monetary policy amplification determined by countercyclical inequality would aggravate the forward guidance puzzle.

The forward guidance puzzle resolution. Below we present the forward guidance experiment as McKay et al. (2016) conduct it in their article.[40] For simplicity, we follow their assumption that the central bank behaves according to an exogenous rule for the real interest rate $r_t = i_t - \overline{m}\mathbb{E}_t\pi_{t+1} = r_t^{nat} + \epsilon_{t,t-j}^{FG}$, and we slightly change the natural rate of interest definition from equation 1. We ignore the effective discount factor shock and simplify the natural rate of interest into $r_t^{nat} = \rho$. We denote the shock term $\epsilon_{t,t-j}^{FG}$ to be the forward guidance shock driven by the central bank announcement about an interest rate change at some time in the future. The time notation indicates the shock to the interest rate that is announced in period $t-j$, but which materializes at time t. Next, the central bank announces that the real interest rate will drop by 1 percentage point for one quarter five years from now, but for the meantime, the interest rate remains at ρ. Therefore, we have $\epsilon_{t+20,t} = -0.01$. Figure 9 below plots the responses of output and inflation to the forward guidance shock. We show a comparison of our model with different variations of the model based on the calibration of the key parameters. The rest of the parameterization is in line with the baseline case presented in section 2.3.

We can observe that the Behavioral Tractable HANK model we use is the one that can provide a solution to the forward guidance puzzle and at the same time incorporate the monetary policy amplification at the time of the shock manifestation.[41]

The THANK model that ignores the cognitive discounting satisfies the stronger reaction at the time when the shock occurs, but actually worsens the forward guidance puzzle. This is caused by $\delta = \left[1 + (\chi-1)\frac{1-s}{1-\lambda\chi}\right]$ for $\chi > 1$ in equation 1. When calibrating $\overline{m} = 1$, there is no term that would sufficiently multiply δ and hence lead to discounting of the future instead of compounding. On the other hand, when we apply only Gabaix (2020) myopic structure, but allow for all households to determine their consumption according to the Euler Equation, the absence of hand-to-mouth agents causes no monetary policy amplification. When we absent both Bilbiie (2024) and Gabaix (2020) characteristics, we end up in the textbook three equation New Keynesian model. Thus, the result is equivalent to the one shown in McKay et al. (2016). Put differently, the

---

[40]We show the properties of the forward guidance resolution within the model in a more elaborate way in our online appendix (link here).

[41]A simple heterogeneous expectations New Keynesian model of Branch and McGough (2009) may also deliver these results as Beqiraj et al. (2019) show. However, this approach fails to deliver the monetary policy amplification on the contemporaneous monetary policy shock.

Figure 9: The forward guidance shock in different models
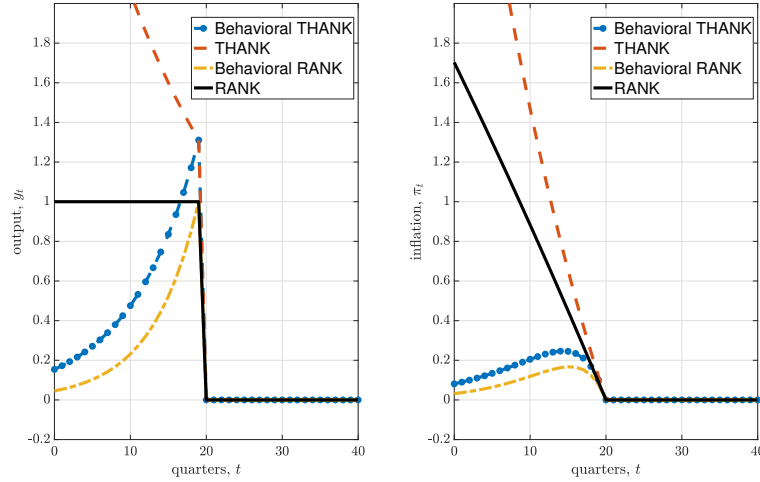
Note: Response on the forward guidance puzzle shock ($\epsilon_{t+20,t}$ = -1 pp) within different versions of the model.

consumption is a step function, as the shock moves the relative price of consumption between the time of the shock and the first period after the shock, but leaves all the periods before and after the shock unaffected. One can see all the mechanisms behind figure 1 if solving equation 1 forward:

$$\tilde{y}_t = - \sum_{j=0}^{\infty} (\overline{m}\delta)^j \frac{1}{\sigma} \frac{(1-\lambda)}{(1-\lambda\chi)} (i_{t+j} - \overline{m}\,\mathbb{E}_{t+j}\,\pi_{t+j+1} - r_{t+j}^{nat}), \tag{15}$$

where we still assume $r_{t+j}^{nat} = \rho$ and $r_{t+j} = i_{t+j} - \overline{m}\mathbb{E}_{t+j}\pi_{t+j+1} = r_{t+j}^{nat} + \epsilon_{t+j,t}^{FG}$. Substituting into r_{t+j}^{nat} yields:

$$\tilde{y}_t = - \sum_{j=0}^{\infty} (\overline{m}\delta)^j \frac{1}{\sigma} \frac{(1-\lambda)}{(1-\lambda\chi)} \epsilon_{t+j,t}^{FG} \tag{16}$$

where the parameters in front of the monetary policy shock may change the results compared to the case of the RANK model.

The indirect general equilibrium effect. The amplification through the indirect effect after a contemporaneous policy rate change is the same as in Bilbiie (2024). Cognitive discounting does not change anything in this respect. To achieve a stronger reaction of output to the policy interest rate change, we must have:

$$\frac{(1-\lambda)}{(1-\lambda\chi)} > 1, \tag{17}$$

which is possible only when $\chi > 1$; i.e., when the consumption inequality between savers (S) and hand-to-mouth (H) defined as $c_t^S - c_t^H = \frac{1-\chi}{1-\lambda}y_t$ is countercyclical (see Bilbiie, 2024).[42]

A.2   Parameterization

Table 1: Baseline parameterization of the model parameters

Parameter	Value	Description
β	0.9925	Subjective discount factor
σ	1	Risk aversion
φ	1	Inverse of the Frisch elasticity
χ	1.48	Inequality determining parameter
λ	0.33	Share of hand-to-mouth agents
s	$\sqrt[4]{0.8}$	Switching parameter in the idiosyncratic shock
ζ_D	0.18	Redistributive parameter
θ	0.8106	Fraction of firms unable to change their prices
$\overline{m}$	0.85	Cognitive discounting
m_π^f	0.85	Firms' cognitive discounting of inflation
m_y^f	0.85	Firms' cognitive discounting of marginal costs
ρ_i	0	Smoothing in the monetary rule
ϕ_π	1.5	Inflation elasticity in the monetary rule
ϕ_y	0.5/4	Output gap elasticity in the monetary rule
ξ	*vary*	Discounting in exponential MA of inflation
ϵ	7.66	Price elasticity of demand
ρ_q	0.75	Persistence of the demand shock
σ_q	0.005	Standard deviation of the demand shock
ρ_u	0.30	Persistence of the supply shock
σ_u	0.0017	Standard deviation of the supply shock

Note: The model baseline parameterization.

---

[42]Evidence for countercyclical inequality given the monetary policy shocks is in Coibion et al. (2017), Samarlna and Nguyen (2019), and Mumtaz and Theophilopoulou (2017). Note that we consider $\lambda < \chi^{-1}$ to rule out the inverse aggregate demand logic shown in Bilbiie (2008).

A.3 The speed or return to the target path

Below we show the IRFs to the natural rate of interest. The shock hits the system in the 4th period and pushes the subjective discount rate below the steady state for 4 periods such that the policy rate reaches ZLB. The shock dissipates after 4 periods with zero persistence. The model used for the exercise is the standard three equation version of NK model from Galí (2015). Hence, the heterogeneous and behavioral features are turned off for simplicity.

The chart displays that the elasticity coefficient in the feedback rule with a makeup rule inside transfers into different speed of return to the target path. We show the case of PLT ($\xi = 0$) with $\phi_{\pi a} = \{1.5, 5, 10, 100\}$ while keeping $\phi_y = 0.125$. Note that $\phi_{\pi a} = 100$ results in the central bank returning the price level back to the target path in 7 periods after the shock arrives (hence 3 after it vanishes). In the case of $\phi_{\pi a} = 10$, the return happens 10 periods after the shock's arrival, $\phi_{\pi a} = 5$ engineers closing the gap after 12 periods, and $\phi_{\pi a} = 1.5$ makeups the previous deviations only after 16 periods.

Figure 10: Different Taylor rule coefficient in PLT driving the speed of return
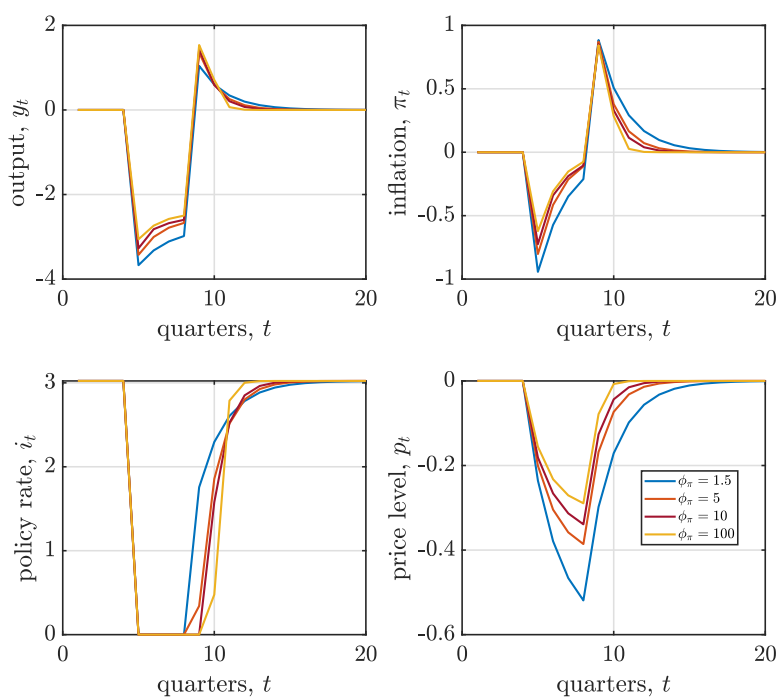
Note: The IRFs to the natural rate of interest comparing the speed of return to the target path by varying the Taylor rule parameter.

A.4   Derivation of the quantitative easing extension

Figure 11: The household side of the model
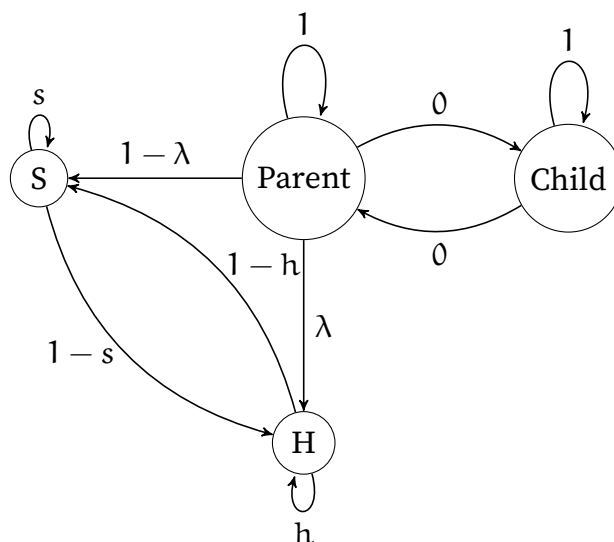
In order to incorporate Sims et al. (2021) into Pfäuti and Seyrich (2023), we now assume that the household side is divided into three agents. Following Sims et al. (2021), we have parents and children household types. However, unlike in Sims et al. (2021), not all the parents can intertemporally move their consumption, because some are hand-to-mouth agents. Hence, the parent section of households is identical to Pfäuti and Seyrich (2023). The only difference is the presence of a regular transfer to the child in the budget constraint. Although the transfer is time-varying, it is not a choice and thus it does not alter the solution of Pfäuti and Seyrich (2023) in any way.[43] This means that the parent structure follows Bilbiie (2024) extended by the subjective discounting of Gabaix (2020) instead of the fully rational expectations operator. This is equivalent to our initial model.

We now also include the child type of household, which does not work and finances consumption by issuing debt and receiving transfers from the parent. We end up with the same setup as in Sims et al. (2021). The only difference is that we use the approach of Gabaix (2020) instead of the rational expectation operator even in this case.

---

[43] We follow Sims et al. (2021) and impose the full-bailout assumption on this transfer. This means that the transfer fully pays off coupon payments plus the outstanding debt that the child accumulated in the past when financing their consumption. For details, see Sims et al. (2021).

The household side after merging Pfäuti and Seyrich (2023) and Sims et al. (2021) can be visualized as in figure 11. While parents and children cannot switch between each other, the unconstrained (S) and constrained (H) households within the parent structure follows a two-state Markov process in line with Bilbiie (2024), where s is the probability of remaining S in the next period and h is the probability of remaining H. The share of H is found as the unconditional probability λ by solving the conventional stationary distribution problem.

Solving the problems of parents and children, we obtain results that are in line with those of Pfäuti and Seyrich (2023) and Sims et al. (2021), respectively. We only extend the latter by the subjective discounting of Gabaix (2020). Hence, the log-linearized Euler equations follow:[44]

$$c_t^P = s\overline{m}\,\mathbb{E}_t\,c_{t+1}^{PS} + (1-s)\overline{m}\,\mathbb{E}_t\,c_{t+1}^{PH} - \psi_c\frac{1}{\sigma}(i_t - \overline{m}\,\mathbb{E}_t\,\pi_{t+1}) \tag{18}$$

$$c_t^B = \overline{m}\,\mathbb{E}_t\,c_{t+1}^B - \frac{1}{\sigma}(i_{t+1}^B - \overline{m}\,\mathbb{E}_t\,\pi_{t+1}), \tag{19}$$

where c^P denotes the consumption of the parents, while c^B is the consumption of the children. Further, i_{t+1}^B is the yield of the long-term bonds that the children issue to finance their consumption.[45]

As in Bilbiie (2024), we can express the consumption of constrained and unconstrained parents as the function of the aggregate product:

$$c_t^{PS} = \frac{(1-\lambda\chi)}{(1-\lambda)}y_t \tag{20}$$

$$c_t^{PH} = \chi y_t \tag{21}$$

The resource constraint of our model is:

$$y_t = (1-z)c_t^P + zc_t^B \tag{22}$$

with:

$$c_t^P = (1-\lambda)c_t^{PS} + \lambda c_t^{PH} \tag{23}$$

---

[44]Our step-by-step derivations do not differ from those of Pfäuti and Seyrich (2023) and Sims et al. (2021).

[45]The whole setting follows Sims et al. (2021). The child has a lower subjective discount factor and the bonds that they issue are conventional decaying coupon bonds. See Sims et al. (2021) for a detailed description.

Substituting 20 and 21 into 18 together with using 22 results in:

$$y_t = \psi_f \, \mathbb{E}_t \, y_{t+1} - (1-z)\psi_c \frac{1}{\sigma}(i_t - \overline{m} \, \mathbb{E}_t \, \pi_{t+1}) - z(\psi_f \, \mathbb{E}_t \, c^B_{t+1} - c^B_t) \tag{24}$$

Following the same steps as Sims et al. (2021) while using the full bailout assumption, leverage constraint, bond market clearing condition, and the central bank's balance sheet yields equation 12:[46]

$$\tilde{y}_t = \psi_f \, \mathbb{E}_t \, \tilde{y}_{t+1} - (1-z)\psi_c \frac{1}{\sigma}(i_t - \overline{m} \, \mathbb{E}_t \, \pi_{t+1} - r^{nat}_t) - z\frac{b^{cb}}{b}(\psi_f \, \mathbb{E}_t \, qe_{t+1} - qe_t),$$

Pertaining to the supply side, note that solving the NK PC from Gabaix (2020) results in:

$$\pi_t = \beta M^f \, \mathbb{E}_t \, \pi_{t+1} + \omega mc_t, \tag{25}$$

where, in our model, due to the absence of technology in the production function, we have $mc_t = w_t$. Assuming that the elasticity parameters in the utility function do not differ between constrained and unconstrained parent households leads to the conventional labour supply equation:

$$\varphi n_t = w_t - \sigma c^P_t, \tag{26}$$

and after using the resource constraint, production function, and market clearing condition, we get:

$$\varphi y_t = w_t - \sigma \frac{(y_t - zc^P_t)}{(1-z)}, \tag{27}$$

where we already know $w_t = mc_t$. Reshuffling it and substituting into 25 gives us the final NK PC as a function of the output gap:[47]

$$\pi_t = \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa^{QE} \tilde{y}_t - \sigma \frac{z}{(1-z)} \frac{b^{cb}}{b} qe_t$$

---

[46]The detailed derivations are attached in the online appendix.

[47]Recall that, in our model the potential output does not change; hence, the fluctuations of the output gap coincide with the output fluctuations.

A.5 Solution techniques

A more comprehensive description of the solution algorithms appears in the online appendix. In the paragraphs below, we outline the algorithms concisely.

The Local Solution. First, we solve the model by local approximation using a first-order perturbation technique. To handle the occasionally binding constraint, we apply the method of Guerrieri and Iacoviello (2015). We use a piecewise linear solution where the first-order approximation to the policy functions around the same steady state is taken for two models. The models are the same type, but in one the OBC is slack, while in the other one it is binding. The resulting solution is non-linear because we obtain a unique set of coefficients for each model. The coefficients are no longer constant, but are time-variant based on the binding or non-binding lower bound. The solution algorithm uses a guess-and-verify method. We guess the periods when the regimes apply, then verify and update the guess if necessary until the guess is verified. A detailed description of the method and its solution algorithm appears in Guerrieri and Iacoviello (2015).

The method cannot capture any precautionary behavior arising from a pure expectation of a binding lower bound in the future, because agents do not pay attention to any information about the path of the future shock. This feature is straightforward when taking into account that Guerrieri and Iacoviello (2015) offer a piecewise linear solution.

The global solution. To take into account a possible uncertainty about a binding lower bound in the future, we also solve the model by deploying the global techniques. We utilize the collocation method, in which we use the finite element approach. Thus, the linear splines are used as the basis functions, considering their convenience in the situation with a kink. Given that the ELB represents the kink in the approximating function, it is suitable to approximate the expectation function instead of the policy function, because the former is a substantially smoother object. Overall, we approximate the functions of the expected inflation and the expected output as linear combinations of the basis functions, such that the approximants satisfy the equilibrium equations at the collocation nodes. We use the Gaussian quadrature scheme (Gauss-

Hermite) to discretize the normally distributed innovations to the natural rate of interest. To find the values of the basis functions coefficients in the collocation equation, we solve the equation as a standard root-finding problem. We apply the *CompEcon* toolkit of Miranda and Fackler (2002) to obtain the approximated solution.

Frantisek Masek
Sapienza University of Rome, Rome, Italy; National Bank of Slovakia, Bratislava, Slovakia; email: frantisek.masek@uniroma1.it

Jan Zemlicka
University of Zurich, Zurich; Switzerland; email: jan.zemlicka@uzh.ch