A Theory of Dynamic Inflation Targets

Christopher Clayton* Andreas Schaab†

June 2021

Abstract

Should central banks' inflation targets remain set in stone? We study a model in which a government designs a mechanism to control inflation setting by a central bank, which suffers from a time consistency problem. The central bank learns about persistent economic shocks affecting optimal inflation, and so influences beliefs of firms who use that information in price setting (Phillips curve). A "dynamic inflation target" implements the constrained efficient inflation level: the central bank reports its target one period in advance, with a linear incentive scheme for deviations from the target. Intuitively, the central bank correctly accounts for its own time consistency problem when setting the target in advance. Both the level and flexibility (slope) of the target can change over time. This mechanism is optimal when the social costs of the incentive scheme are negligible relative to the inflation-output trade-off.

*Clayton: Yale School of Management. Email: christopher.clayton@yale.edu

†Schaab: Toulouse School of Economics. Email: andreas.schaab@tse-fr.eu

We are grateful to Fabrice Collard, Emmanuel Farhi, Xavier Gabaix, Sam Hanson, Christian Hellwig, Matteo Maggiori, Jeremy Stein, Ludwig Straub, and seminar participants at Harvard, TSE, and Yale SOM for helpful comments and suggestions.

1 Introduction

Since their inception in the early 1990s, many central banks' inflation targets have evolved substantially. For example, the Bank of New Zealand has announced at least four major updates to its target definition since 1990.[1] Similarly, the Bank of Canada undergoes regular reviews of its inflation target at 5-year intervals. More recently, the U.S. Federal Reserve engaged in a long-term strategic review of its operating framework. This review was partly motivated by the persistent decline in the natural rate of interest and the accompanying concern of an increase in the incidence of future ZLB spells.[2] Overall, central banks have exercised substantial discretion over inflation target adjustments over this period.

In academic discourse, an important motivation for inflation targets is the interaction between a time consistency problem and central bank private information (e.g. Barro and Gordon (1983)).[3] Prior work has shown that a static inflation targeting mechanism can resolve the resulting inflationary bias in either fully static environments or when shocks are uncorrelated across time (e.g. Walsh (1995), Athey et al. (2005)). These frameworks motivate inflation targets as desirable mechanisms but do not speak to the empirical regularity that central banks regularly update their targets. In practice, central banks often motivate such target adjustments with reference to structural economic change, which presupposes that shocks are correlated across time.

Our paper studies a dynamic monetary policy game in the presence of persistent shocks, which introduce a motive for target adjustment. As in previous work, the central bank is faced with a time consistency problem that arises because firms' inflation expectations affect economic activity contemporaneously ("Phillips Curve"). When shocks are persistent, however, central bank private information also becomes persistent, which gives rise to novel informational frictions that affect the monetary policy game. In particular, firms form expectations about future inflation based on their beliefs about the current and future state of the economy. When the central bank has private information about persistent shocks, firms learn from the central bank's actions and update their beleifs about the distribution of future shocks. In this setting, the central bank has an additional

[1]See for example McDermott and Williams (2018). The Bank of New Zealand's initial target postulated an inflation band of 0-2%. The band was revised in 1996 to 0-3% and again in 2002 to 1-3%. Another revision in 2012 added an explicit focus on the 2% target midpoint.

[2]See for example Clarida (2019). In August 2020, the Fed concluded its review by adopting a target that aims to "achieve inflation that averages 2% over time" (Powell (2020)). At the same time, commentators have also suggested an explicit upward revision in the inflation target, for example Blanchard et al. (2010), Ball (2013) and Krugman (2014).

[3]There is much empirical support for the existence of central bank private information. For example, see Romer and Romer (2000), Kuttner (2001), Gürkaynak et al. (2005), Campbell et al. (2012), Krishnamurthy and Vissing-Jorgensen (2012), and Lucca and Moench (2015) among many others.

incentive to distort firm beliefs through the revelation of private information in order to induce a more favorable contemporaneous inflation-output trade-off. In our paper, the persistence of shocks therefore interacts substantively with the two core ingredients of the benchmark monetary policy game framework, asymmetric information and time consistency problems.

In this setting, we study the problem of a government that designs a "transfer" mechanism to control central bank behavior. These transfers can be penalties that hurt both the government and central bank, or incentive payments that benefit the central bank at the expense of the government. The use of penalties in particular finds much empirical support in the institutional frameworks that govern central bank behavior in many countries. In practice, such penalties take the form of legislative scrutiny, reputational risk, the structure of terms of office, and even the threat of termination. In the U.S., for example, the central bank Chairwoman is directly accountable to Congress in the form of bi-annual and extraordinary Congressional testimonies, and her term must be renewed on a four year basis.[4]

The main result of this paper is that a *dynamic inflation target* implements the constrained efficient level of inflation. A dynamic inflation target assigns a penalty (or transfer) based on deviations of inflation from the inflation target, where the target is always equal to next-period expected inflation. Critically, both the target level as well as penalties imposed for deviations from the target, i.e. the target's flexibility, can be updated *one period in advance* of the central bank's inflation choice. This means that in each period the central bank both sets inflation based on the existing target and updates the target for the next period if necessary. Our mechanism recognizes that, in the presence of persistent structural change, society benefits when the central bank has the flexibility to adjust its target. At the same time, our result highlights that some commitment in the target adjustment process is necessary to align central bank incentives. The dynamic inflation target is also an optimal mechanism whenever the social cost of the incentive mechanism governing central bank behavior is negligible relative to the inflation-output trade-off.

We begin our analysis in Section 2 with a simple version of our model based on a log-linearized New Keynesian model at a distorted steady state, so that expected future inflation lowers current output through the standard New Keynesian Phillips Curve. The economy transitions stochastically between two persistent states, which differ in terms of the implied welfare gains from stimulating output. We show that the central bank can

[4]Among many others, Rogoff (1985), Walsh (1995), Svensson (2010), and Halac and Yared (2019) provide additional examples and further discussion of penalties that can be and are imposed on central banks in practice.

achieve the constrained efficient inflation policy by adopting a rigid inflation target which it can update one period in advance with discretion. The rigid inflation target prescribes an arbitrarily large penalty for deviations from the target, and so affords the central bank no flexibility in its inflation choice within a given period. Our illustrative model deliberately shuts off this within-period dimension of the standard commitment-flexibility trade-off.

The central bank's ability to update the target means that there is flexibility across periods to respond to persistent shocks to the optimal inflation level. At the same time, our result highlights that a certain degree of commitment in this target adjustment process is necessary to achieve constrained efficiency: When the central bank is required to make target adjustments at least one period in advance, it internalizes its own future time consistency problem. In the presence of persistent shocks, therefore, a new and inter-temporal trade-off between commitment and flexibility emerges. Society benefits when the central bank has the flexibility to adjust its target as the economy undergoes structural change, but a control mechanism is necessary to guide target adjustment in order to guard against an un-anchoring of inflation expectations. Our results suggest that requiring adjustment in advance is a desirable control mechanism.

To further illustrate the importance of this target adjustment process, we study central bank behavior both in the case of pure discretion (no target), and in the case where the central bank uses a target but has discretion over when it updates this target. Equilibrium inflation under these two cases coincides, meaning that an unconstrained target adjustment process does no better than pure discretion. Intuitively, a central bank with discretion over target adjustments can simply update its target at the beginning of each period to coincide with the desired level of inflation under pure discretion. Firms in the previous period then correctly anticipate that the central bank will update its target, which un-anchors firms' inflation expectations and undermines the inflation target.

We present our full model in Section 3. The model features a continuum of economic states which evolve over time with persistence, as well as more general social preferences and inflation-output relationships, under which next period inflation expectations affect current period output. The government designs a transfer mechanism to incentivize how monetary policy is set by the central bank, accounting not only for the time consistency problem of the central bank, but also for the information friction that arises because firms learn about the state of the economy from the choices of the central bank. We begin by characterizing the *full-information constrained efficient* allocation under which the government, central bank, and firms all observe the economic shock at each date, and the government mandates a path of inflation under commitment. This commitment solution under full information provides a benchmark for allocative efficiency in the economy.

In Section 4, we present the main result of this paper: A dynamic inflation target implements the constrained efficient level of inflation. Moreover, it is also an optimal mechanism when the social cost of the transfer mechanism governing central bank behavior is negligible relative to welfare losses that arise from the inflation-output trade-off. In particular, the dynamic inflation target disciplines the central bank's inflation choice in each period by providing it with linear transfers for deviations of inflation from the target level, which is always equal to expected inflation. Importantly, our mechanism delegates to the central bank the authority to adjust both the *slope* (i.e. the marginal inflation penalty, or target "flexibility") and the *level* (or "intercept") of the linear inflation target, but it can only do so one period in advance of its inflation choice. We furthermore show that implementing the dynamic inflation target requires minimal information: the current target's slope and level are a sufficient statistic for the entire past history of shocks when evaluating how the target should be updated.

The intuition behind the optimality of the dynamic inflation target is as follows. First, the slope of the target manages the time consistency problem in the central bank's inflation choice by penalizing inflation in excess of the target level. Because current output depends on next-period inflation expectations, the appropriate penalty function is linear in inflation.[5] Second, requiring the target to be adjusted one period in advance allows the central bank to internalize its own future time consistency problem. As in the simple two-state model of Section 2, this constraint provides the necessary commitment in the target adjustment process to align central bank incentives. As a result, inflation expectations remain anchored to the target level even though it can be adjusted by the central bank across time.

Third, the dynamic inflation target resolves the information friction, under which the central bank has an incentive to bias firm expectations in pursuit of a more favorable inflation-output trade-off. Suppose that a central bank considers misreporting its type and, in doing so, biasing firm inflation expectations downward. This would boost current period output due to the Phillips Curve relationship, thus benefiting the central bank. At the same time, however, the dynamic inflation target requires that the target level is always equal to firm inflation expectations. It would therefore also fall when the central bank misreports, implying higher future penalties for inflation in excess of the now lower target level. At the constrained efficient allocation, these effects exactly offset each other, giving no benefit to the central bank from misreporting its type to manipulate firm beliefs. The mechanism therefore resolves the information friction that arises from shock persistence.

The optimality of a dynamic inflation target in the presence of persistent shocks sug-

[5] This logic is the same as in the static model of Walsh (1995).

gests that controlled target adjustment may be preferable to a perpetually fixed target. Our paper presents a simple mechanism that can help inform implementation in practice. The key idea of practical relevance that emerges in this paper is that, while target adjustment can be fully delegated to the central bank, it is important to ensure that targets are not updated contemporaneously. In practice, our mechanism resembles closely the institutional process adopted by the Bank of Canada which undergoes target reviews at fixed five-year intervals. Our results underscore that a controlled target adjustment process can provide flexibility in response to persistent structural change without risking an un-anchoring of inflation expectations. Lastly, our paper emphasizes that target adjustments in response to persistent shocks can involve changes in both the level and the flexibility of the inflation target.

In Section 5, we present two applications of our framework motivated by empirical evidence and policy debates on structural change in the U.S. that is relevant for monetary policy. First, we consider a setting in which the slope of the Phillips Curve is subject to persistent shocks that may alter the welfare implications of stimulating output. Second, we capture in reduced form the potential welfare implications of a persistent decline in the natural rate of interest which pushes the economy towards an Effective Lower Bound on nominal interest rates, which we model as a change in the socially desired rate of inflation. These applications are illustrative in part because they result in sharp limiting cases of our dynamic inflation target mechanism. When the slope of the Phillips Curve is subject to persistent shocks, the optimal mechanism features a time-varying target slope while the level of the inflation target remains constant. In the context of persistent shocks to the socially desired rate of inflation, on the other hand, the level of the inflation target becomes time-varying while the slope is constant. In both cases, we derive sharp characterizations of potential welfare gains under a dynamic inflation target relative to alternative mechanisms.

We finally consider two important extensions of our analysis in Section 6. First, we revisit the argument of Rogoff (1985) that installing a conservative central banker who places a larger subjective penalty on inflation can replicate the benefits of a formal inflation target. In our setting, a conservative central banker fails to achieve the efficient allocation when the dynamic inflation target prescribes a time-varying target slope. However, a transition process whereby more (less) conservative central bankers are appointed whenever the dynamic inflation target's slope increases (decreases) could in principle achieve the efficient outcome.

Second, we allow for transfers to the central bank to be costly from the government's perspective, so that transfer costs must be accounted for when designing an optimal mechanism. Although a dynamic inflation target still implements the constrained efficient allocation, the government no longer finds it exactly optimal to do so due to costs of the transfer scheme. Nevertheless, we show that the properties of the optimal mechanism still bear features resembling a dynamic inflation target, including that the optimal mechanism reverts to a dynamic inflation target at the extremes of the shock distribution.

While monetary policy is the primary focus of this paper, our results could be applied more broadly to principal-agent settings where "moving goal posts" are desirable due to a combination of persistent private information and time consistency problems arising through expectations. For example, the sovereign debt literature commonly features a time consistency problem that arises because future exchange rate devaluations affect current bond prices. Our model could be applied without change to a problem in which a government issues a fixed (nominal) face value of short-term debt every period and also chooses the exchange rate. Under this interpretation, the expected future exchange rate affects current period consumption through the price of short-term debt. Persistent shocks affect the value of consumption to the representative household, but are private information of the domestic government and so not known by foreign lenders.

Related literature. The paper most closely related to ours is Halac and Yared (2014). Their paper uses dynamic mechanism design techniques to study the effects of persistent private information in the fiscal policy context in a delegation framework, with time inconsistency driven by present bias due to quasi-hyperbolic discounting. By contrast, we study persistent private information in the monetary policy context with transfers, with a time consistency problem resulting from inflation expectations in a Phillips Curve relationship. The combination of persistent private information and time constiency via inflation expectations gives rise to a novel set of informational effects present in our model. These effects occur because firms form inflation expectations based on their beliefs about the current (persistent) state, which they infer from the report of the central bank under the mechanism. More broadly, the literature on the commitment versus flexibility trade-off with quasi-hyperbolic agents studies both transfer mechanisms (DellaVigna and Malmendier (2004), Galperti (2015), Beshears et al. (2020)) and delegation mechanisms (Amador et al. (2006), Halac and Yared (2018)) to control these agents.[6]

We also relate to the literature on time consistency and the commitment flexibility trade-off in monetary policy.[7] This literature approaches this question both using transfer

[6] More generally, Pavan et al. (2014) provide necessary and sufficient conditions for implementability in a general principal-agent framework with transfers and persistent shocks, and uses them to characterize optimal mechanisms.

[7] For example, see Kydland and Prescott (1977), Barro and Gordon (1983), Canzoneri (1985), Rogoff (1985),

or penalty (i.e. separable money burning) based mechanisms, and delegation mechanisms. In the former approach, Walsh (1995) shows that an inflation target is an optimal mechanism in a static context with transferable utility. The linear form of their static target follows the same intuition as the within-period linear form of our dynamic target. Halac and Yared (2019) use a framework with socially costly penalties that can be imposed on the central bank to study the trade-off between instrument-based rules and target-based rules. In the delegation context, Athey et al. (2005) considers a dynamic monetary policy framework with independent shocks. They show that the optimal mechanism is static, and characterized by bounds on inflation, with full flexibility between the bounds. Waki et al. (2018) considers a related framework with independent shocks in which the optimal mechanism consists of history dependent bounds on inflation. Relative to this literature, we study inflation setting with transfers (or penalties) and persistent private information, which gives rise to novel information frictions due to inflation expectations in a Phillips Curve type relationship. Similar to Walsh (1995) and Halac and Yared (2019), we adopt an approach using transfers and penalties rather than the delegation approach.

Our discussion also connects to a long literature that studies the optimal rate of inflation.[8] More recently, many papers have investigated quantitatively whether a fall in the natural rate of interest in the presence of a zero lower bound constraint could quantitatively justify higher inflation targets.[9] In our paper, we take as given that persistent structural shocks can alter the welfare implications of inflation and, consequently, the socially desired rate of inflation. We ask if and how a central bank should respond to such shocks in the presence of persistent private information and time consistency problems.

2 Two State Model

We begin with a simple and illustrative version of our main model, which we use to illustrate the key results and intuitions. We present the full model in Section 3.

Our economy has a government, a monetary authority (central bank), and a continuum of small firms. Time is infinite and discrete, indexed by $t = 0, 1, \ldots$. Allocations in this economy are summarized by two scalar variables, inflation π_t and output y_t. Both inflation

Cukierman and Meltzer (1986), Persson and Tabellini (1993), Svensson (1995) and Walsh (1995) for the former among many others. More broadly, there has been a long tradition considering the implications of private information for the design of policy. For example, see Backus and Driffill (1985), Sleet (2001), and Angeletos et al. (2006) among many others.

[8] For example, see Schmitt-Grohé and Uribe (2010) and references therein.

[9] For example, see Coibion et al. (2012), Kiley and Roberts (2017), Andrade et al. (2018) and Eggertsson et al. (2019) among others.

and output lie in closed sets, $\pi_t \in [\underline{\pi}, \overline{\pi}] \subset \mathbb{R}$ and $y_t \in [\underline{y}, \overline{y}] \subset \mathbb{R}$.

Both the government and central bank in our economy share the same social preferences, given by

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t \left[-\frac{1}{2}\alpha\pi_t^2 + \frac{1}{\theta_t} y_t \right], \tag{1}$$

where β is the discount factor, and $\theta_t \in \{\underline{\theta}, \overline{\theta}\}$ denotes an economic shock that affects the gains from stimulating output.[10] Equation (1) is drawn from a New Keynesian loss function at a distorted steady state. The shock θ_t follows a two-state Markov process with transition probability λ. The government does not directly observe this shock and must therefore delegate inflation policy to the central bank, which is tasked with identifying θ_t and then setting monetary policy.

The relationship between inflation and output is determined by price-setting firms. It is a log-linearized New Keynesian Phillips Curve, given by

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa y_t, \tag{2}$$

where the date t conditional expectation of future inflation depends both on firm beliefs about the current state θ_t (due to persistence) as well as firm beliefs about how future monetary policy will be determined.

To simplify analysis, in this section we take the limiting case where $\alpha = 0$, and hence there is no direct inflation penalty. However, there is still a penalty for inflation operating through the Phillips Curve. In particular, internalizing the Phillips Curve (2) into social preferences (1), we obtain the social welfare function

$$\mathbb{E}\frac{1}{\kappa}\left[\frac{1}{\theta_0}\pi_0 + \sum_{t=1}^{\infty} \beta^t \left[\frac{1}{\theta_t}\pi_t - \frac{1}{\theta_{t-1}}\mathbb{E}_{t-1}\pi_t\right]\right], \tag{3}$$

which both the government and central bank seek to maximize. Notice that θ_t acts equivalently in the welfare function to a change in the slope of the Phillips Curve. In other words, the effective Phillips Curve slope at date t is $\kappa\theta_t$.

[10] As will become clear below, the shock θ_t we study in this illustrative benchmark corresponds to a persistent change in the slope of the Phillips curve. A vibrant discussion has emerged recently focused on the apparent fall in the slope of the Phillips curve since the Great Recession. In Section 5, we study in more detail two applications of the general model we present in Section 3: The first application follows the present interpretation of θ_t as a persistent change in the slop of the Phillips curve. In our second application, we study in reduced-form the implications of the sectoral decline in the natural rate of interest for monetary policy.

2.1 Constrained Efficient and Discretionary Monetary Policy

We begin by characterizing a constrained efficient allocation that could be achieved if the government and firms observed θ_t in each period, and the government mandated *with commitment* the inflation decisions of the central bank. In other words, the government mandates at date 0 an inflation decision $\pi_t(\theta^t)$ to be set along every history θ^t. This provides an efficiency benchmark which respects the Phillips Curve relationship between inflation and output determined by firms. This full-information constrained efficient allocation follows immediately from optimizing the government objective in equation (3).

Proposition 1. *A full-information constrained efficient allocation is* $\pi_0 = \overline{\pi}$ *and*

$$\pi_t = \begin{cases} \overline{\pi}, & \text{if } \theta_{t-1} = \overline{\theta} \\ \underline{\pi}, & \text{if } \theta_{t-1} = \underline{\theta} \end{cases} \qquad \forall t \geq 1$$

All proofs can be found in Appendix A. An optimal allocation for inflation is governed by the Philips Curve inflation-output trade-off. If $\theta_{t-1} = \overline{\theta}$, then the cost of future inflation through the Phillips Curve is low, that is $\frac{1}{\theta_{t-1}}$ is low. As a result, the marginal benefit $\frac{1}{\theta_t}$ of inflation tomorrow is at least as large as the marginal cost $\frac{1}{\theta_{t-1}}$ of inflation today, regardless of the realization $\theta_t \in \{\underline{\theta}, \overline{\theta}\}$. It is thus (weakly) optimal to set inflation to be as high as possible *next period*, $\pi_t = \overline{\pi}$, regardless of the realization of θ_t. Notice that in this sense inflation is "backward looking," in that date t inflation is determined entirely by the date $t-1$ shock, that is by the impact on date $t-1$ output. If on the other hand $\theta_{t-1} = \underline{\theta}$, then the cost of future inflation through the Phillips Curve is high, that is $\frac{1}{\theta_{t-1}}$ is high. As a result, the marginal cost of future inflation weakly exceeds the marginal benefit regardless of the realization of θ_t, and it is (weakly) optimal to set inflation to be as low as possible, $\pi_t = \underline{\pi}$. As such, one optimal allocation can be entirely determined by the state at the prior date, which represents the degree of the time consistency problem.[11]

Discretionary monetary policy. Proposition 1 supposes that a government observes θ_t and mandates inflation policy with commitment. Consider alternatively the case where the central bank observes θ_t as private information. Inflation policy is delegated to the central bank, which has full discretion over monetary policy. We obtain the following result.

[11] Of course, there are other optimal allocations that feature any choice $\pi_t \in [\underline{\pi}, \overline{\pi}]$ when $\theta_t = \theta_{t-1}$. We have chosen the allocation that will have a simple decentralization, as we show below.

Proposition 2. *Under discretion, an equilibrium central bank allocation is $\pi_t = \overline{\pi}$ for all t and θ_t.*

In any period t, the central bank finds it optimal to maximize flow utility state by state. In particular, the central bank neglects the impact of inflation on the previous period's Phillips Curve under discretion, which no longer serves as a constraint on the problem. This results in inflationary bias and reflects a standard Barro and Gordon (1983) time consistency problem. This time consistency problem motivates studying a mechanism to control the central bank's inflation policy.

2.2 Implementing Constrained Efficiency by Changing Inflation Targets

We now propose a method of implementing the constrained efficient allocation via changing inflation targets. This benchmark serves as a preview of our main result in the full model. In particular, we define a *rigid inflation target* to be an inflation level π_t^* that the central bank must hit exactly at date t. That is, the central bank is required to set $\pi_t = \pi_t^*$. This is a strong inflation target that offers no flexibility in adjusting inflation at date t.

We now show how a central bank, choosing its own rigid inflation target, can achieve the constrained efficient allocation, provided that it updates its target *one period in advance.*

Proposition 3. *Suppose that in each period $t-1$, the central bank chooses a rigid inflation target $\pi_t^* \in [\underline{\pi}, \overline{\pi}]$ that applies in the next period and requires $\pi_t = \pi_t^*$. Then, the central bank chooses*

$$\pi_t^* = \begin{cases} \overline{\pi}, & \theta_{t-1} = \overline{\theta} \\ \underline{\pi}, & \theta_{t-1} = \underline{\theta} \end{cases} \qquad \forall t \geq 1$$

and so implements a full-information constrained efficient allocation.

Proposition 3 shows that a central bank with the ability to update its inflation target can achieve a constrained efficient inflation policy, provided that the central bank updates its target one period in advance. Intuitively, by setting the inflation target in period $t-1$ the central bank internalizes the impact of date t inflation on the date $t-1$ Phillips curve. In setting its rigid target, the central bank is in effect setting its own inflation policy for the next period. As a result, the central bank agrees with the government's commitment solution about the optimal policy for the *next* period, even though it disagrees with the commitment solution about the optimal inflation policy for the *current* period. The central bank therefore chooses the correct target for the next period, while the rigid target set in the prior period forces it to adopt the commitment policy for the current period.

In the environment of Proposition 3, there is no further role the government plays once the targeting framework is established. The central bank has full flexibility to update its target in advance every period, and may set its target to whatever level it wishes.

Naive Target Adjustment. To illustrate the importance of adjustment one period in advance, let us consider instead a benchmark of *naive target adjustment*: the central bank has full flexibility to update its target π_t^* contemporaneously at the beginning of date t, before setting inflation for date t.

Proposition 4. *Under naive target adjustment, an equilibrium central bank target and allocation is $\pi_t^* = \pi_t = \overline{\pi}$ for all t and θ_t.*

Whereas Proposition 3 shows that a rigid target that can be updated one period in advance achieves a constrained efficient allocation, Proposition 4 shows that a rigid target that can be updated at any time leads to reversion to monetary policy under discretion. The inflationary bias which the inflation target is supposed to correct re-emerges, even though the target is rigid and allows for no deviations from the target. If the central bank has the authority to update its inflation target contemporaneously under discretion, then it can simply set the target to equal the level of inflation in the optimal allocation under discretion. Therefore, even though firm expectations are anchored to π_t^*, firms anticipate in period $t-1$ that the central bank will update the target to $\pi_t^* = \overline{\pi}$. Naive target adjustment thus undermines the target, and de-anchors expectations from the desired commitment solution.

Discussion. The results of this section highlight that when shocks are persistent, there is in principle a two-layered trade-off between commitment and flexibility, not only in the central bank's inflation choice but also in its adjustments of the inflation target itself.

In our framework, a standard Barro and Gordon (1983) time consistency problem results in inflationary bias, but the constrained efficient inflation policy is determined solely by the shock in the previous period. This means that within a given period, there is no benefit to flexibility and hence no commitment-flexibility trade-off in inflation choice. As a result, a rigid inflation target that mandates an inflation level implements the correctly inflation policy for any given period.

Although constrained efficient inflation is constant within period, it is not constant between periods. In particular, it changes in response to persistent shocks, for example as the economy undergoes structural change. A new trade-off between commitment and flexibility emerges: society benefits when the central bank has the flexibility to adjust its operating framework in response to persistent shocks (i.e. changing the rigid target). Under a complete lack of commitment, however, inflationary bias reemerges and inflation expectations become un-anchored because, under fully discretionary adjustment, the central bank adjusts its target contemporaneously to match its desired inflation level.

Our main result shows that when the rigid target is set one period in advance, this intertemporal commitment-flexibility trade-off is fully resolved. Our mechanism delegates to the central bank the authority to adjust its own target in response to persistent shocks. And since it can only do so one period in advance, the central bank in fact internalizes its own future time consistency problem and adjusts the target correctly.

3 Model

We now present our full model. As in Section 2, the economy has a government, a monetary authority (central bank), and a continuum of small firms. Time is infinite and discrete, indexed by $t = 0, 1, \ldots$, with allocations summarized by inflation $\pi_t \in [\underline{\pi}, \overline{\pi}] \subset \mathbb{R}$ and output $y_t \in [\underline{y}, \overline{y}] \subset \mathbb{R}$. The relationship between inflation and output will again be determined by price-setting firms and, therefore, a Phillips Curve.

The government and central bank interact in a principal-agent framework. The central bank (agent) learns about the state of the economy, and uses this information to determine monetary policy. However, the central bank is subject to a time-consistency problem in the tradition of Kydland and Prescott (1977) and Barro and Gordon (1983). As a result, the government (principal) designs a mechanism to determine how the bank sets monetary policy, subject to truthful reporting conditions. Firms are not directly under the control of the government, so their actions (i.e. the determination of output) will be taken as a constraint on the problem.

Government. The social preferences of the government are given by

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t \mathcal{U}_t(\pi_t, y_t, \theta_t), \tag{4}$$

where β is the discount factor, and $\theta_t \in \Theta = [\underline{\theta}, \overline{\theta}]$ denotes an economic shock with conditional density $f(\theta_t|\theta_{t-1})$.[12] The government does not directly observe the shock θ_t, which

[12] Persson and Tabellini (1993) also consider a general social welfare function where shocks enter that change the welfare consequences inflation.

will be observed and reported by the central bank.

Interpretation of θ_t. We think of θ_t as corresponding to the true economic shock to the economy, which the central bank learns each period. In this sense, the government welfare function reflects a notion of social welfare that incorporates the true shock – for example, a true supply side shock or a shock to the loss function for inflation and output deviations. However, both the firms and the government rely on the central bank to collect and report information about that true type.[13] As such, both the government and firms form decision rules based on the *reported* type, not the true type. The government's beliefs matter because it designs the mechanism that maps the reported type (beliefs) into an allocation rule. Firms' beliefs also matter because firms form a decision rule based on those beliefs. In both cases, that decision rule responds to beliefs, while the "true" welfare function is related to the true type.

In Section 5, we develop two applications of our theory, choosing particular shocks θ_t that are motivated by empirical evidence and recent policy debates on important structural changes.

Central Bank. The central bank has preferences over both social welfare, and transfers T_t from the government, so that its preferences are[14]

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t \Big[\mathcal{U}_t(\pi_t, y_t, \theta_t) + T_t \Big]. \tag{5}$$

The central bank observes the shock realization θ_t in every period, and so is tasked with choosing inflation and output in each period, subject to output determination by firms as

[13] There is a long tradition in macroeconomics to motivate and study monetary policy games when the central bank has private information. For example, see Sargent and Wallace (1975), Barro and Gordon (1983), Canzoneri (1985), Rogoff (1985), Walsh (1995), Athey et al. (2005) and many others. There exists much empirical support for the existence of central bank private information. Romer and Romer (2000) show that the difference between the Federal Reserve's private inflation forecasts and commercial inflation forecasts is a significant predictor of commercial forecast errors. Lucca and Moench (2015) document sizable excess returns on U.S. equities leading up to scheduled Federal Open Market Committee (FOMC) meetings. These findings suggest the existence of substantial (private) information content in FOMC announcements. Krishnamurthy and Vissing-Jorgensen (2012) find strong empirical support for a signaling channel of unconventional monetary policy, whereby large-scale asset purchases between 2009 and 2012 worked to a large extent by conveying private information to financial market participants. Similarly, work by Kuttner (2001) and Gürkaynak et al. (2005) shows that Federal Reserve announcements are associated with significant price effects that are not due to changes in the policy rate itself. Building on this work, Campbell et al. (2012) show that asset prices and commercial macroeconomic forecasts respond strongly to the information content in FOMC announcements, even when conventional monetary policy is constrained by the Effective Lower Bound on nominal interest rates.

[14] Walsh (1995) considers a transfer contract in a static setting.

defined below.

In writing preferences, we have assumed that T_t is welfare neutral from the perspective of the government, and is only used as a control mechanism. This can be seen as a limiting case where the costs of controlling central bank behavior are negligible relative to the underlying social welfare problem. In Section 6, we explicitly study the case where transfers are not neutral from the perspective of the government. However, it should be noted that our main result in Section 4 is an implementability result that holds whether or not transfers are neutral from the government's perspective.

Interpretation of Transfers. Rather than monetary transfers, the practical analogs of the control mechanism T_t may be closer to policies such as Congressional scrutiny, reputational risk, or firing the central banker.[15] For example, a central bank being awarded high T_t may face a low degree of Congressional scrutiny in its policy determination. So far, T_t could be a utility transfer, or it could be a form of money burning. Of course, one implementation of T_t could be to use explicit monetary incentives.

Firms and Phillips Curve. Firms, whose actions are not directly controlled by the government or central bank, create a "Phillips Curve" relationship between inflation and output, given by

$$y_t = F_t(\pi_t, \mathbb{E}_t[\pi_{t+1}|\mu_t]). \tag{6}$$

The object μ_t is the (posterior) belief of firms about the current state θ_t. Beliefs are formed based on the report of the central bank. Since θ_t features persistence, beliefs about the current state of θ_t affect beliefs about next period's state θ_{t+1} and hence next period's monetary policy and inflation choice π_{t+1}. As firms cannot be directly controlled, equation (6) is an implementability condition from the perspective of the government and central bank.

To simplify exposition, we internalize the Phillips Curve relationship of equation (6) into preferences, and write

$$U_t(\pi_t, \mathbb{E}_t[\pi_{t+1}|\mu_t], \theta_t) = \mathcal{U}_t(\pi_t, F_t(\pi_t, \mathbb{E}_t[\pi_{t+1}|\mu_t]), \theta_t)$$

where we use U_t for more compact notation.

[15] In the U.S., for example, this process is multifaceted. The central bank Chairwoman is directly held accountable to Congress in the form of bi-annual, as well as extraordinary, Congressional testimonies. The institutions adopted by modern central banks to allow for active monitoring by stakeholders and mantain accountability vary across countries but most are highly complex.

Lucas Critique. A key concern of this Phillips Curve relationship is a Lucas critique – firms' price-setting behavior may change in response to changes in the monetary policy regime, such as target changes.[16] Our Phillips Curve relationship is robust to a Lucas critique provided that expected future (next period) inflation is sufficient for determining how changes in future policies affect firm behavior. For example, higher expected inflation may lead firms to increase the frequency with which they update prices, altering the slope of the Phillips Curve.

3.1 Constrained Efficient Allocation

Before characterizing the mechanism structure, we characterize the constrained efficient allocation that could be achieved if the government and firms observed θ_t, and the government mandated with commitment the inflation decisions of the central bank. Given full information, firms' posterior beliefs are the degenerate distribution which places all mass on θ_t, which we denote by $\mu_t = \theta_t$. This provides an efficiency benchmark which respects the Phillips Curve relationship between inflation and output determined by firms. This is the analog of Proposition 1 in Section 2, and follows from optimizing equation (5) after internalizing the Phillips curve relationship (6).

Proposition 5. *The full-information constrained efficient allocation of the government is given by*

$$\frac{\partial U_t}{\partial \pi_t} = -\frac{1}{\beta} \frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}(\pi_t \mid \theta_{t-1})} \quad \forall t \geq 1 \tag{7}$$

$$0 = \frac{\partial U_0}{\partial \pi_0} \tag{8}$$

Because there is no Phillips Curve constraint for π_0 (it would occur at date -1), the optimality condition for π_0 simply sets the marginal value of increasing inflation to 0, yielding a standard first-order condition that would also be chosen by the central bank under discretion. For $t \geq 1$, the first-order condition reflects the Phillips Curve relationship, and internalizes that inflation at date t affects output at date $t-1$. The left-hand side (LHS) of equation (7) is date t-adapted, whereas the right-hand side (RHS) is date $t-1$-adapted. Therefore, the RHS is constant from the perspective of time t, implying that the marginal (flow) utility from inflation is constant at date t in histories θ^t proceeding from the same history θ^{t-1}. Although optimal inflation is no longer fully determined by the previous

[16] See e.g. L'Huillier and Schoenle (2019).

period's shock, as it was in Section 2, the time consistency problem reflected on the RHS is still summarized by information available at date $t-1$.

Discretionary monetary policy. Suppose that we maintain full information, but the central bank were left to its own devices to set inflation with discretion, that is $T_t = 0$. At period t for any t, the central bank finds it optimal to set $\partial U_t / \partial \pi_t = 0$ state by state. In particular at date t, the central bank neglects the impact of inflation on the previous period's Phillips Curve, which no longer serves as a constraint of the problem. This results in inflationary bias and reflects a standard Barro and Gordon (1983) time consistency problem. This motivates studying a mechanism to control the central bank's inflation policy.

A sufficient statistic. It will be useful to define the wedge

$$\nu_{t-1} = -\frac{1}{\beta} \frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}(\pi_t \mid \theta_{t-1})} \tag{9}$$

evaluated at the constrained efficient allocation. This wedge ν_{t-1} is a date $t-1$ adapted constant that is a sufficient statistic for the shock history θ^{t-1} in determining the allocation rule $\pi_t, \pi_{t+1}, ...$ for inflation. In other words, the constrained efficient allocation from dates t and onward can be calculated with the knowledge of the wedge ν_{t-1}, without knowing the exact shock history θ^{t-1} that gave rise to it. ν_{t-1} is the wedge between the constrained efficient allocation rule for π_t and the central bank's allocation rule under discretion. In other words, this wedge reflects the inflationary bias that arises from the central bank's time consistency problem.

3.2 Mechanism Structure

We study direct and *full-transparency* mechanisms, under which the central bank makes a report of its type each period.[17] By full transparency, we mean that there is no pooling of central bank types in reporting. We denote this report by $\tilde{\theta}_t$. Along the equilibrium path, agents' posterior will therefore be the degenerate distribution at the reported type, or $\mu_t = \tilde{\theta}_t$. Note that we abuse notation here because μ_t is a full distribution in general.

Restricting attention to full transparency mechanisms is not without loss of generality. In principle, the government could want to pool central bank types to shroud the type and manipulate firm beliefs. In assuming mechanisms under which the central bank truthfully

[17] We study direct mechanisms, invoking the revelation principle.

reveals its type, we assume away such motivations. Given that central bank transparency has become an increasingly prominent focal point over the last two decades, we impose the restriction to full transparency.[18]

A mechanism in our model is a mapping from the history of *reported* types into a transfer and allocation rule, given by $(\pi_t, T_t) : \Theta^t \to \mathbb{R}^2$. Although the date t allocations depend on the entire history of reported types, we will show state space reduction results which allow us to characterize sufficient statistics for information histories.

3.3 Incentives, Time Consistency, and Information

At every date t, the central bank makes a report $\tilde{\theta}_t$ of its true type θ_t. We use a first order approach to incentive compatibility in deriving results.[19] To understand the forces underlying the incentives of the central bank in reporting its type, we study the local incentives of a central bank to misreport its type along a path where all past and future selves are reporting their types truthfully. In doing so, we identify three driving forces of the model: a time consistency problem from the Phillips Curve, and two informational problems. These informational problems arise due to shock persistence.

Suppose that we define the value function of the central bank by $\mathcal{W}_t\left(\theta^t\right)$, where θ^{t-1} is the history of reported types whereas θ_t is always the current true type. The central bank's reporting problem at date t, when all past and future selves are reporting truthfully, is given by

$$\mathcal{W}_t\left(\theta^t\right) = \max_{\tilde{\theta}_t} \mathbb{E}_t\left[\sum_{s=t}^{\infty} \beta^s \left(U_s(\pi_s, \mathbb{E}_s[\pi_{s+1}|\tilde{\theta}_s], \theta_s) + T_s\right) \mid \theta_t\right]$$

where the history-dependence is implicit in the policies (π_s, T_s). The current report $\tilde{\theta}_t$ determines not only the current allocation (π_t, T_t) but also all future allocations due to history dependence. Furthermore, the report $\tilde{\theta}_t$ also affects firm and government beliefs $\tilde{\theta}_t$ about the current state. Formally, the marginal change in welfare to the central bank from misreporting its type at date t is given by

$$\underbrace{\frac{\partial U_t}{\partial \pi_t}\frac{\partial \pi_t}{\partial \tilde{\theta}} + \frac{\partial T_t}{\partial \tilde{\theta}}}_{\text{Inflation and Incentives}} + \underbrace{\frac{\partial U_t}{\partial \mathbb{E}_t}\mathbb{E}_t\left[\frac{\partial \pi_{t+1}}{\partial \tilde{\theta}} + \pi_{t+1}\frac{\partial f(\theta_{t+1}|\tilde{\theta})/\partial \tilde{\theta}}{f(\theta_{t+1}|\tilde{\theta})}\right]}_{\text{Phillips Curve and Firm Expectations}} + \underbrace{\beta \mathbb{E}_t\left[\frac{\partial \mathcal{W}_{t+1}}{\partial \tilde{\theta}}|\theta_t\right]}_{\text{Future Allocations}}$$

[18] See e.g. Powell (2019).

[19] We make use of techniques in Pavan et al. (2014) for a first order approach with persistent shocks.

There are three key forces that appear in this decision rule: one reflecting time consistency, and two reflecting information considerations related to persistent shocks.

The first, reflecting time consistency, is marked by the *absence* of any terms that capture the impact on the previous period's Phillips Curve. This reflects the Barro and Gordon (1983) time consistency logic, that the central bank neglects the previous period's Phillips Curve when determining current inflation. This is the first force that the government must account for when designing the mechanism.

The second force, related to persistent shocks, is that the central bank evaluates expectations with respect to the *true* type θ_t, whereas the government designs policy and forms expectations based on the *reported* types. This creates an informational friction where the central bank has an incentive to manipulate the beliefs of the government, while its own beliefs are held constant.

A third and related force is that an analogous information manipulation incentive exists for firm beliefs. Firms form beliefs based on the reported type, meaning that the central bank can manipulate their actions with the information it reports. The inflation expectations that appear in the Phillips Curve are formed based on firms' beliefs and, therefore, central bank reports. Concretely, this force is captured by the derivative of the density f in the reported type. In the conventional New Keynesian framework, the central bank will want to bias firm inflation expectations *downward* in order to improve the contemporaneous inflation-output trade-off.

These three forces affect not only the central bank's inflation decision, but also its reporting decision. When designing a mechanism, the government must account for all three of these effects.

3.4 Naive Target Adjustment

It is well known that an inflation target is a method of controlling inflationary bias, which is reflected in our two state model in Proposition 3. This proposal arises in static mechanism design problems as a way to enforce efficient inflation policies. Relative to the static setting, however, we feature a dynamic problem of persistent information, where the optimal inflation rate moves substantially over time, and where there are additional informational incentives over the beliefs of firms and the government. As a result, a static target would not be able to achieve the constrained efficient level of inflation, and is unlikely to constitute an optimal mechanism.

However, motivated by the previous literature, a benchmark proposal would be to simply grant the central bank flexibility to adjust its target. This would allow the central bank to accommodate structural shocks that change the inflation-output trade-off in a fundamental way. It is clear, however, that this type of mechanism would simply reintroduce the Barro and Gordon (1983) time consistency problem and undermine the target. The central bank would have an incentive in every period to evaluate its optimal policy $\frac{\partial u_t}{\partial \pi_t} = 0$, and then to set its inflation target to coincide with that inflation level. Rather than only resetting its target in response to structural shocks, the central bank also resets its target to accommodate its inflationary bias.

We can see this formally by adopting the rigid inflation target of the two-state model, from which we obtain a version of Proposition 4.

Proposition 6. *Under full information and flexible adjustment of a rigid inflation target, the central bank adjusts the target π_t^* at the beginning of date t to satisfy*

$$\frac{\partial U_t(\pi_t^*, \mathbb{E}_t[\pi_{t+1}|\theta_t], \theta_t)}{\partial \pi_t^*} = 0$$

and so implements the policy under discretion.

The intuition of Proposition 6 follows as in Proposition 4. Recognizing that the central bank would set its inflation target each period to coincide with its discretionary policy, firms in the previous period form inflation expectations that coincide with expected inflation under discretion. In other words, firm inflation expectations are no longer anchored to the current target because they know it will simply be adjusted in the next period to accommodate the inflationary bias. This results in a de-anchoring of inflation expectations. Naive adjustment undermines the target by allowing the central bank to set its target to its desired inflation, rather than the intended goal of forcing inflation to match the target.

4 Dynamic Inflation Target

In this section, we show the main result of our paper, which is that a "dynamic inflation target" mechanism can implement the constrained efficient allocation when the target is set by the central bank *one period in advance*. This is the analog of Proposition 3 in Section 2.

Equations (7) and (9), along with their sufficient statistic implications, suggest a mechanism that uses the transfer rule T_t to penalize deviations from a target. An inflation target of this form seeks to correct the time consistency problem in the central bank's inflation policies by incentivizing it to set inflation close to the target. However, in the presence of persistent structural shocks, the target might need to be adjusted over time to accommodate a changing efficient level of inflation. Due to structural shocks, the optimal inflation rate may drift far from the present target in a persistent manner, implying large potential gains from letting the target adjust. The commitment-flexibility trade-off that motivated the inflation target in the first place may itself be subject to structural change. Indeed, this is precisely reflected in the time variation of the full-information constrained efficient allocation in the presence of persistent θ_t shocks.

We look over a class of mechanisms defined by the transfer rule

$$T_t = -b_{t-1}(\pi_t - \tau_{t-1}), \tag{10}$$

where τ_{t-1} is the target and b_{t-1} is the slope of the punishment for inflation in excess of the target.[20] Under our proposed mechanism, the parameters (b_{t-1}, τ_{t-1}) of the transfer rule T_t are specified one period in advance, in that they are based only on information knwon at date $t-1$. As a result, when the central bank reports its type at date t, two things happen. First, its type report maps into a contemporaneous inflation policy π_t, which in turn generates a transfer T_t based on the target parameters (b_{t-1}, τ_{t-1}) specified in the previous period. Second, the report also maps into target parameters (b_t, τ_t) for the transfer rule in the next period. In sum, the mechanism is a mapping $(\pi_t, b_t, \tau_t) : \Theta^t \to \mathbb{R}^3$ from the history of reported types into inflation for the current period and the target for the next period. Then, the transfer is determined from the current inflation level π_t and the current target (b_{t-1}, τ_{t-1}). In practice, we can also think of the central bank as directly choosing inflation and its own future target, represented by (π_t, b_t, τ_t), from among the set of triples that follow from the same history θ^{t-1} of reported types prior to date t.

Our main result is that this dynamic inflation target implements the constrained efficient level of inflation in a locally incentive compatible mechanism, which moreover admits a key state space reduction property.

Proposition 7 (Dynamic Inflation Targeting)**.** *A dynamic inflation target implements the constrained efficient allocation in a locally incentive compatible mechanism. The target* $\tau_{t-1} = \mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]$ *and the intercept* $b_{t-1} = \nu_{t-1}$ *are given by their constrained efficient values (Proposition 5). The target* (τ_{t-1}, b_{t-1}) *is a sufficient statistic at date t for the history* $\tilde{\theta}^{t-1}$ *of reported types.*

Proposition 7 shows that in a setting with persistent structural shocks, the constrained efficient allocation can be implemented by a simple dynamic inflation target. Under

[20] The structure of our target mechanism and, in particular, its linearity are similar to Walsh (1995) in a static setting.

the dynamic inflation target, inflation always meets the target in expectation, that is $\tau_{t-1} = \mathbb{E}_{t-1}\pi_t$. The slope of the target mechanism is always given by $b_{t-1} = \nu_{t-1}$. The inflation target prescribed by our mechanism is dynamic in the sense that both τ_t and b_t are time-varying.

At each date t, the target inherited from the previous period corrects the time consistency problem in the central bank's contemporaneous inflation choice. The intuition behind the form of the target is analogous to the well-known logic from the static setting (e.g. Walsh (1995)). In particular, the time consistency problem being corrected by the inflation target is that expected date t inflation, $\mathbb{E}_{t-1}\pi_t$, appears in the Phillips Curve at date $t-1$. However, at date t this is no longer a constraint of the problem, and so is neglected by the central bank. By applying a penalty $\nu_{t-1}(\pi_t - \mathbb{E}_{t-1}\pi_t)$ to inflation exceeding expectations, the central bank overcomes the time consistency problem. To understand why this penalty is linear, the Phillips Curve relationship in equation (6) depends on expected inflation, meaning that $\frac{\partial y_{t-1}}{\partial \pi_t} = \frac{\partial y_{t-1}}{\partial \mathbb{E}_{t-1}\pi_t} f(\theta_t|\theta_{t-1})$. This means that the probability-normalized impact of future inflation on prior period output is the same across all future states θ_t proceeding from the same history θ_{t-1}. This generates the linear form of the inflation penalty.

However, the inflation target also needs to be adjusted in response to structural shocks in order to allow inflation to coincide with the constrained efficient allocation. There are two key insights of Proposition 7. The first is that the central bank optimally resets this target one period in advance. That is, when the central bank observes a persistent shift in the constrained efficient inflation level, it adjusts its inflation target for the *next* period in response to this shift. However, the current target remains in effect for the current period inflation policy. The second is that both components of the target, the intercept τ_{t-1} and the slope b_{t-1}, are subject to being changed when the target is updated. The slope and intercept reflect two different economic facets of the dynamic inflation target. The intercept reflects the "level" of the target, that is value that the central bank is expected to hit on average. The slope, on the other hand, reflects the "flexibility" of the target, that is how sharp the punishments are when the central bank exceeds its target level. An increase in τ_{t-1} means that the target level for the next period has increased, which reduces central bank penalties for high levels of inflation. By contrast, an increase in b_{t-1} means that the target flexibility has been reduced, which increases central bank penalties for high levels of inflation.

Moreover, the current target $(\nu_{t-1}, \mathbb{E}_{t-1}\pi_t)$ is a sufficient statistic for the entire history θ^{t-1} of shock realizations. This sufficient statistic property follows precisely as in the constrained efficient allocation, as the slope of the target ν_{t-1} summarizes the time consistency problem that arises from the previous period's Phillips Curve. This implies that all history dependence of the mechanism is captured in the inflation target itself. This greatly reduces the knowledge required for the central bank to adjust its target. At date t, the central bank needs to know what its current target is, but not the history under which that target arose.

The dynamic inflation target of Proposition 7 is successful in implementing the constrained efficient outcome because it is able to overcome all three of the frictions we identified in Section 3: time consistency, government belief manipulation, and firm belief manipulation.

First, the current transfer rule T_t, based on the current target (b_{t-1}, τ_{t-1}), corrects the current period's time consistency problem by imposing a marginal penalty $b_{t-1} = \nu_{t-1}$ on current inflation. Given that ν_{t-1} is precisely the wedge between the allocation under discretion and the constrained efficient allocation, this aligns the preferences of the central bank at date t with that of the government operating with commitment. Furthermore, when adjusting the target intercept for next period, b_t, the central bank internalizes the current period's Phillips Curve relationship, which features expected inflation for the next period. As a result, as in Proposition 3, the central bank also internalizes its own future time consistency problem and corrects it.[21]

If shocks were not persistent, time consistency would be the only force in the model, and Proposition 7 would still characterize optimal policy. This reflects the logic of the static model, with the additional insight that by setting the target in advance, the time consistency problem in target adjustment is also overcome. In this sense, the optimality of the dynamic inflation target directly reflects the logic of the standard time consistency problem. However, if there is no shock persistence and shocks are small, there may be little value to adjusting the target over time as the resulting commitment versus flexibility trade-off may be inconsequential.[22]

By contrast when shocks are large and persistent, there may be a strong incentive to adjust the target over time. At the same time, however, the two additional incentive problems associated with shock persistence emerge: government and firm belief manipulation. It is therefore surprising that this simple target and adjustment mechanism remains relevant. The dynamic inflation target remains relevant because at the constrained efficient allocation, these two informational effects exactly offset one another. In particular, when the central bank considers on the margin changing its report $\tilde{\theta}_t$, this generates a welfare

[21] This is also similar to the static setting, where the central bank is willing "ex ante" to set up a targeting mechanism for itself. It is also closely related to the literature on optimal mechanisms to control present bias (e.g. Amador et al. (2006)), where agents are willing to set up mechanisms to control their own time consistency problems.

[22] It should be noted that even with iid shocks, the target may in principle move around over time, which follows from noting that a price level target is generally an optimal commitment policy to one-time shocks in an infinite horizon.1

impact related to beliefs which is given by[23]

$$\underbrace{\frac{\partial U_t}{\partial \mathbb{E}_t[\pi_{t+1} \mid \tilde{\theta}_t]} \frac{d\mathbb{E}_t[\pi_{t+1} \mid \tilde{\theta}_t]}{d\tilde{\theta}_t}}_{\text{Firm Beliefs}} + \underbrace{\beta \nu_t \frac{d\mathbb{E}_t[\pi_{t+1} | \tilde{\theta}_t]}{d\tilde{\theta}_t}}_{\text{Government Beliefs}}$$

Under the constrained efficient allocation, we have $\frac{\partial U_t}{\partial E(\pi_{t+1}|\tilde{\theta}_t)} + \beta \nu_t = 0$, and hence these two terms cancel out. In economic terms, on the one hand the central bank wishes to bias inflation expectations of the government *upward*, which increases the target and reduces expected penalties for exceeding its target. On the other hand, the central bank also wishes to bias the inflation expectations of firms *downward*, in order to economize on the Phillips Curve relationship and improve the contemporaneous inflation-output trade-off. At the constrained efficient allocation where the slope of the target is exactly equal to the Phillips Curve impact, the marginal reduction in penalties for exceeding the target is exactly equal to the Phillips Curve impact, and these two forces exactly offset one another.

It is worth noting that these two forces exactly offset each other because central bank and government preferences align at date $t-1$ over the future inflation-output trade-off. However, although the central bank at date $t-1$ agrees with the government about optimal inflation from dates t and onward, internalizing the Phillips Curve relationship, it also wishes to manipulate firm beliefs. Recognizing this, the government designs its transfer rule to offset firm belief manipulation. As a result, the two informational effects exactly offset each other.

Dynamic inflation target as an optimal mechanism. Proposition 7 makes a statement about implementability: a dynamic inflation target *can* implement the constrained efficient allocation of inflation. It does not state that it is an optimal mechanism. It is immediate, however, that if we assume the control mechanism has no direct welfare consequence to the government – that is, T_t does not appear in the government objective function – then the dynamic inflation target is also an optimal mechanism, because it maximizes the government's objective function. This implies that if the government's primary concern is with the inflation-output trade-off, and not with the costs of implementing the mechanism, the dynamic inflation target is optimal.

On the nature of inflation targets. The dynamic inflation target of Proposition 7 sheds light on the important features of the target in anchoring inflation expectations. In par-

[23]See the proof of Proposition 7 for a detailed derivation.

ticular, the target's efficacy does not result from it being constant and unchanging, but rather from it being committed to sufficiently in advance. De-anchoring of expectations arises in the naive adjustment process because the central bank is able to match its target to its desired inflation contemporaneously, rather than forcing inflation to match the target. By ensuring the target is set sufficiently in advance, the central bank is forced to set its inflation to its target instead. This reflects the same fundamental insight that underlies Proposition 3.

Two-state model and dynamic inflation targets. In Section 2, we provided a variant of Proposition 7 that used a rigid inflation target, which mandated an inflation level. We can also implement the constrained efficient allocation with a dynamic inflation target. In particular, suppose the central bank can choose between a high target, $(b_t, \tau_t) = (\frac{1}{\overline{\theta}}, \overline{\pi})$, and a low target, $(b_t, \tau_t) = (\frac{1}{\underline{\theta}}, \underline{\pi})$. If a central bank chooses a high target, it will always find it (weakly) optimal to pick $\pi_t = \overline{\pi}$ in the next period, regardless of the shock realization. Similarly if the central bank chooses a low target, it will always find it (weakly) optimal to pick $\pi_t = \underline{\pi}$ the next period. The dynamic inflation target therefore provides another implementation of the constrained efficient allocation.

In the two-state model, the level and flexibility of the dynamic inflation target move in opposite directions: as the headline rate increases, the flexibility also increases, in the sense that the slope of punishment falls. This happens because optimal inflation in the two-state model is directly pegged to the time consistency problem and Phillips Curve relationship. When the Phillips Curve impact on previous period output is strong, target flexibility and optimal inflation are both low. Similarly, if the Phillips Curve impact on previous period output is small, target flexibility and optimal inflation are both high.

4.1 Evolution of the Target

The characterization of the target in Proposition 7 can be used to understand the evolution of the target. Combining the first order condition of the constrained efficient allocation (7) with the definition of the updated target ν_t, and making the natural assumption that

$\frac{\partial F_t/\partial \pi_t}{\partial F_t/\partial E_t \pi_{t+1}} < 0$ local to the constrained efficient allocation,[24] we obtain a law of motion,

$$\Delta \nu_t = \underbrace{-\frac{\partial U_t}{\partial \pi_t}}_{\text{Excess Inflation}} + \underbrace{\left(1 - \left|\frac{\partial F_t/\partial \pi_t}{\partial F_t/\partial \mathbb{E}_t \pi_{t+1}} \beta\right|\right) \nu_{t-1}}_{\text{Time Consistency}}, \tag{11}$$

reflecting the extent to which the slope of the target varies over time.

The first effect, "excess inflation," indicates that the slope of the target *increases* as the central bank reduces marginal utility from inflation. In conventional models, this corresponds to higher levels of inflation. In the absence of a persistent structural shock, this term reflects a conventional New Keynesian Phillips Curve (NKPC) logic. When the central bank generates excess inflation in order to stimulate output, the central bank also has a motivation to *reduce* future inflation in order to capitalize on the current inflation-output trade-off. This leads the central bank to increase the slope of the target, punishing future inflation exceeding the target and reducing inflation expectations.

The second force, "time consistency," reflects the relative effects of current and future inflation on current output through the Phillips Curve. In the standard NKPC, $\left|\frac{\partial F_t/\partial \pi_t}{\partial F_t/\partial \mathbb{E}_t \pi_{t+1}} \beta\right| = 1$, and this term drops out. More generally, when the relative effect of current inflation is large relative to future inflation, that is $\left|\frac{\partial F_t/\partial \pi_t}{\partial F_t/\partial \mathbb{E}_t \pi_{t+1}} \beta\right| > 1$, the time consistency problem is small relative to demand stabilization, leading to a reduction in the slope of the target and allowing for greater future inflation. By contrast when $\left|\frac{\partial F_t/\partial \pi_t}{\partial F_t/\partial \mathbb{E}_t \pi_{t+1}} \beta\right| < 1$, the time consistency problem is relatively severe, and the slope of the target is increased, pushing down future inflation.

Finally, we can consider how the target intercept changes in response to a structural shock. Considering a marginal increase in the structural shock, we have

$$\frac{d\tau_{t-1}}{d\theta_{t-1}} = \underbrace{E_{t-1}\left[\pi_t \frac{\partial f(\theta_t|\theta_{t-1})/\partial \theta_{t-1}}{f(\theta_t|\theta_{t-1})}\right]}_{\text{Expectations}} + \underbrace{\frac{\partial \nu_{t-1}}{\partial \theta_{t-1}} \mathbb{E}_{t-1}\left[\frac{\partial \pi_t}{\partial \nu_{t-1}}\middle|\theta_{t-1}\right]}_{\text{Target Slope Adjustment}}. \tag{12}$$

There are two effects of the structural shock on the target. The first effect, "expectations," reflects that the probability measure over future states changes in response to the shock. If a higher θ_t raises the probability of high-inflation states, then the target intercept τ_{t-1} increases as well. The expectations effect therefore implies that, in response to persistent shocks, the target intercept can change even when the target slope remains constant.

[24] This assumption implies that current and future inflation move current output in opposite directions, as for example in the standard log-linearized New Keynesian Phillips Curve.

If shocks were fully transitory, on the other hand, the probability measure would not be affected and no adjustment in the target intercept would be required. The second effect, "target slope adjustment," reflects the extent to which a change in the slope of the target impacts optimal future inflation. In the natural case where $\frac{\partial \pi_t}{\partial \nu_{t-1}} < 0$, the upward adjustment of the target intercept is *amplified* (*dampened*) when the target slope is reduced (increased). In economic terms, if a structural shock leads to a decrease in the slope ν_{t-1}, then the central bank will find it optimal to generate higher average levels of inflation in the next period, since the penalty for exceeding the target has been reduced. Firms anticipate this, so that inflation expectations increase for a given state θ_{t-1} and associated conditional density f. As a result, the target level τ_{t-1}, which is set equal to firm inflation expectations, also increases.

In sum, when the economy experiences a structural shock θ_t, both components of the target may be affected. The intercept of the target is directly affected by changes in expectations, but is also affected indirectly if the structural shock leads to a change in target slope. The target slope updates as a result of the structural shock if it leads to a fundamental change in either the motivation of the central bank to generate excess inflation, or if it alters the nature of the time consistency problem.

4.2 Welfare Gains from a Dynamic Inflation Target

We now look to understand what the possible welfare gains might be from instituting a dynamic inflation target, relative to a permanent and static target. Suppose that instead of a dynamic target, the central bank instead adopted a static target (ν^*, τ^*). The following proposition describes the first order welfare gains (in terms of allocative efficiency) from moving from the static target to a dynamic inflation target.

Proposition 8. *To first order, the welfare gains in allocative efficiency from moving from a static target (ν^*, τ^*) to the dynamic inflation target (ν_{t-1}, τ_{t-1}) of Proposition 7 are*

$$\mathbb{E} \sum_{t=1}^{\infty} \beta^t \left[\nu_{t-1}^* - \nu^* \right] \left[\mathbb{E}_{t-1} \pi_t^* - \tau_{t-1} \right]$$

The first order welfare gains available from moving to a dynamic inflation target depend on two forces. The first, $\nu_{t-1}^* - \nu^*$, is the intertemporal variation in the time consistency problem under the static target (where ν_{t-1}^* is the time consistency wedge evaluated at the allocation obtained under the static target). When $\nu_{t-1}^* > \nu^*$, the time consistency problem is more severe than the slope imposed ν^*, and inflation is excessively high. The second term, $\mathbb{E}_{t-1}\pi_t^* - \tau_{t-1}$, is the difference between inflation expectations under the static target and inflation expectations under the dynamic target. High welfare gains are therefore available when a large excess time consistency problem, $\nu_{t-1}^* - \nu^*$, coincides with substantial excess inflation, $\mathbb{E}_{t-1}\pi_t^* - \tau_{t-1}$, relative to the constrained efficient inflation level. The dynamic inflation target thus allows welfare gains not only by allowing for greater inflation when the static target would be too severe, but also by allowing for lower inflation when the static target would be too flexible.

4.3 Target Adjustment in Practice

Proposition 7 provides a mechanism to implement the constrained efficient allocation, which takes the form of a dynamic inflation target. The key property of the target adjustment process is that it occurs one period in advance.

The inflation target adjustment in our model happens at the frequency of the underlying structural shock, in the sense that large movements in the target will generally coincide with large and persistent shock realizations. This suggests that, in practice, our mechanism either requires that a new target is announced sufficiently far in advance of taking effect, or that the target is adjusted contemporaneously but for a long enough interval of time. This latter method minimizes the time consistency problem by forcing the target to remain in force for a duration of time, so that the central bank is guided primarily by the long-term consequences of the target and not by short-term inflationary bias.

What does *a period in advance* mean in practice? In the context of our mechanism, the relevant notion of period length is tied to the underlying time consistency problem and thus to the degree of price stickiness. The inflation target must be updated a period in advance so that the central bank properly internalizes the inflation expectations term in the Phillips Curve. This term emerges because firms adjust prices only infrequently. Average price duration would therefore be a key empirical moment to inform an implementation of our mechanism in practice. Target changes should be announced sufficiently far in advance that most firms will have had a chance to reset prices before going into effect.

A related force in practice which we do not model explicitly in this paper is the transmission lag of monetary policy.[25] As this lag increases, the effective frequency of

[25] Monetary policy transmission is known to have "long and variable lags" (Friedman (1972)). See also Rudebusch and Svensson (1999), Goodhart (2001) and Taylor and Wieland (2012) among others, as well as Havranek and Rusnak (2012) for a summary of the literature. Many central banks have adopted policy horizons ranging between 12 to 24 months.

price resetting also increases from the perspective of time consistency. A longer lag would reduce the interval of time constituting a period.

On the other hand, it is also important to know the duration of structural shifts in the underlying shock θ_t. When the frequency of structural shocks is low relative to the frequency of price adjustment, a central bank observing a structural shock does not expect to observe another one until well after firm prices have adjusted. We would therefore expect lower frequency structural shocks to increase the notion of period length.

The notion of the "correct" length of a period must consider all three of these forces. Fundamentally, the trade-off reflects a form of commitment versus flexibility trade-off in target adjustment. On the one hand, granting greater flexibility (more frequent target adjustments) allows responses to structural shocks. On the other hand, greater flexibility also exacerbates the time consistency problem.

Comparing our mechanism to real-world inflation target regimes. A close analogue to our mechanism in practice is the adjustment process for the Bank of Canada.[26] The Bank of Canada revisits its inflation target every five years at fixed intervals, with the possibility of adjusting its target for the next five-year interval based on new information. Our results suggest that this type of adjustment process can implement the efficient level of inflation by providing a means of responding to structural economic shocks, without generating substantial time consistency problems when updating the target.

5 Applications

In this section, we study the results of Section 4 in our two main applications. Our first application is a model in which the slope of the Phillips Curve is subject to structural change over time. In our second application, we model in reduced form the implications of a persistent decline in the natural rate of interest, which pushes the economy closer to the effective lower bound on interest rates, and thus motivates a higher constrained efficient level of inflation. In the first application, the flexibility (slope) of the target will change over time, while the level (intercept) will remain fixed. By contrast in the second application, the level will change over time, while the flexibility will remain constant.

[26]For background, see https://www.bankofcanada.ca/core-functions/monetary-policy/inflation/.

5.1 Changing Slope of the Phillips Curve

In our first application, we build on the two-state model studied in Section 2. In this application, θ_t represents a persistent shock to the social benefit of stimulating output. As before, θ_t can therefore also be interpreted as a direct shock to the slope of the Phillips curve.[27]

Social welfare is given by a New Keynesian loss function at a distorted steady state,

$$\mathcal{U}_t(\pi_t, y_t, \theta_t) = -\frac{1}{2}\left[\pi_t^2 + \gamma y_t^2\right] + \theta_t y_t,$$

where we set $\gamma = 0$ to obtain closed form solutions. The input-output relationship of firms is given by a New Keynesian Phillips Curve,

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa y_t.$$

Finally, we assume that shocks are fully persistent, with $\mathbb{E}_t \theta_{t+1} = \theta_t$.

Given this specification, we obtain the following characterization of inflation policy and the dynamic inflation target, which are derived from Propositions 5 and 7.

Proposition 9. *The constrained efficient inflation level is*

$$\pi_t = \frac{\theta_t - \theta_{t-1}}{\kappa} \tag{13}$$

The dynamic inflation target that implements the constrained efficient allocation is

$$b_{t-1} = \frac{1}{\kappa}\theta_{t-1} \tag{14}$$

$$\tau_{t-1} = 0 \tag{15}$$

With shocks to the slope of the Phillips Curve, the dynamic inflation target features movement in the target slope, but not in the intercept. In particular, an increase in θ_{t-1} increases the target slope, reducing the flexibility of the target and increasing the penalty for exceeding the fixed target. The increase in the target slope arises because a higher desire to stimulate output today means that the impact of future inflation on current output

[27]The changing slope of the U.S. Phillips curve has garnered much attention since the Great Recession. See for example Blanchard (2016), Gal\'\i and Gambetti (2019), Rubbo (Rubbo), and Del Negro et al. (2020) among others. This debate on the recent flattening of the Phillips curve has also engulfed monetary policy discourse in recent years. See for example Brainard (2015).

through the Phillips Curve is also more pronounced. This exacerbates the time consistency problem in the next period, and leads to a greater penalty for future inflation.

However, the headline rate itself does not structurally shift. This is because shocks are fully persistent, and so the expected value of simulating output next period, $\mathbb{E}_t\theta_{t+1}$, is always equal to the current value of stimulating output, θ_t. Note that although the target is fixed over time, actual inflation deviates from the target following a change in θ. In particular, when $\theta_t > \theta_{t-1}$, then the value of stimulating output at date t is unexpectedly high relative to the value of stimulating output at date $t-1$, and so positive inflation is allowed.

Welfare gains. Suppose instead that the central bank was forced to use a static inflation target (ν^*, τ^*). We can sharply characterize the welfare gains achieved under a dynamic inflation target in this application.

With a static target, optimal inflation is $\pi_t^* = \frac{\theta_t}{\kappa} - \nu^*$. The first order approximation to welfare losses has $\nu_{t-1}^* = \frac{1}{\kappa}\theta_{t-1}$ and $\mathbb{E}_{t-1}\pi_t^* = \frac{\theta_{t-1}}{\kappa} - \nu^*$. Therefore, welfare gains from shifting to a dynamic inflation target are characterized, to first order, by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\nu_{t-1}^* - \nu^*\right]\left[\mathbb{E}_{t-1}\pi_t^* - \tau_{t-1}\right] = \frac{1}{\kappa^2}\sum_{t=1}^{\infty}\beta^t\left[\mathrm{var}_0(\theta_{t-1}) + \left(\mathbb{E}_0\theta_{t-1} - \kappa\nu^*\right)^2\right]. \quad (16)$$

If we furthermore choose ν^* to minimize the first order welfare gains, then $\nu^* = \frac{1}{\kappa}\mathbb{E}_0\theta_{t-1}$, that is ν^* is set to the average time consistency problem. In this case, we have first order welfare gains of

$$\frac{1}{\kappa^2}\sum_{t=1}^{\infty}\beta^t \mathrm{var}_0(\theta_{t-1}).$$

This highlights that the first order feasible welfare gains are related to the conditional variance of future economic shocks. In a model with small and idiosyncratic shocks, this conditional variance will be small, and welfare gains from a dynamic inflation target will be small. However in a model with large and persistent shocks, this conditional variance can be large.

Partial shock persistence. In Proposition 9, we focused on the case of full shock persistence. We can also consider partial shock persistence, where $\mathbb{E}_{t-1}\theta_t = (1-\rho)\tilde{\theta} + \rho\theta_{t-1}$ for $0 \le \rho \le 1$ and where $\tilde{\theta} \in [\underline{\theta}, \overline{\theta}]$ is a constant. The solution for inflation π_t remains the same, as does the slope b_{t-1}. However, the dynamic inflation target now features an intercept that is characterized by

$$\tau_{t-1} = \mathbb{E}_{t-1}\pi_t = \frac{(1-\rho)(\tilde{\theta} - \theta_{t-1})}{\kappa}.$$

When $0 \leq \rho < 1$, the target intercept decreases in θ_{t-1}, while the slope b_{t-1} increases. In other words, a high (low) shock θ_{t-1} is associated with a decrease (increase) in both the level and flexibility of the inflation target. This is because when shocks are partially but not fully persistent, a high shock is expected to revert partially towards $\tilde{\theta}$ in the next period, meaning that the value of stimulating output next period is expected to be lower than the value of stimulating current output. This leads to lower desired future inflation and larger penalties for future inflation that exceeds the target.

Suppose that we go all the way to the case of iid shocks, so that $\rho = 0$ and $\theta_t = \tilde{\theta} + \epsilon_t$ and $\text{var}_0(\epsilon_t) = \sigma_\epsilon^2$. In this case, we have $\tau_{t-1} = \frac{-1}{\kappa}\epsilon_{t-1}$ and $b_{t-1} = \frac{1}{\kappa}\tilde{\theta} + \frac{1}{\kappa}\epsilon_{t-1}$. Notice that in this case, we have $\text{var}_0(\tau_{t-1}) = \text{var}_0(b_{t-1}) = (\frac{\sigma_\epsilon}{\kappa})^2$. Thus if the variance of the iid shock is small, then the variance of the dynamic inflation target will also be small.

5.2 Implications of Declining Natural Rate of Interest

A vibrant debate has recently emerged on the question of how monetary policy should respond to the observed decline in the natural rate of interest (e.g. Laubach and Williams (2016)). In the presence of an Effective Lower Bound (ELB) on interest rates, a decline in the natural rate implies that nominal interest rates are closer to the ELB on average and there is less room for central banks to use conventional monetary policy during recessions. Many observers in the U.S. have hence advocated for an increase in the Federal Reserve's inflation target.[28]

A reduced-form representation of this argument is that as the natural rate persistently falls, all else equal, the socially desired level of inflation rises so that nominal interest rates are higher and further from the ELB constraint. We capture this intuition by writing the social welfare function as

$$\mathcal{U}_t = -\frac{1}{2}\alpha(\pi_t - \theta_t)^2 + y_t,$$

where θ_t can now be interpreted as a shock to the socially desirable level of inflation. Inflation evolves according to a standard New Keynesian Phillips Curve

$$\pi_t = \beta\mathbb{E}_t\pi_{t+1} + \kappa y_t.$$

[28]Concluding its longer-term strategic review in August 2020, the Fed adopted an *average inflation target*.

From here, we obtain the following characterization of the constrained efficient allocation and the dynamic inflation target.

Proposition 10. *The constrained efficient inflation level is*

$$\pi_t = \theta_t \tag{17}$$

The dynamic inflation target that implements the constrained efficient allocation is

$$b_{t-1} = \frac{1}{\kappa} \tag{18}$$

$$\tau_{t-1} = \mathbb{E}_{t-1}\theta_t. \tag{19}$$

With shocks to the socially desirable level of inflation, constrained efficient inflation takes on a simple form, with $\pi_t = \theta_t$ not depending on the previous period's shock. However, this does *not* imply that there is no time consistency problem, since the dynamic inflation target has a slope $b_{t-1} = \frac{1}{\kappa}$, related to the slope of the Phillips Curve. In this case, because welfare is linear in output and the government and firms have the same discount factor, the marginal benefit of using inflation today to stimulate output, $\frac{1}{\kappa}$, exactly equals the marginal cost of higher expected inflation in the previous period Phillips Curve, which was $\frac{1}{\kappa}\frac{\beta}{\kappa} = \frac{1}{\kappa}$ after discounting appropriately. As a result, the constrained efficient allocation simply sets the loss from inflation to zero, that is $\pi_t = \theta_t$. However, this is not the same as discretionary monetary policy, which would set $\pi_t > \theta_t$, accounting for the benefit of stimulating current output but neglecting the effect on previous period output.

As a result, movements in the desired level of inflation generate movement in the target intercept, but not in the target slope. Target adjustment moves the level of the target, but not its flexibility. This is in stark contrast to our first application and Proposition 9 where, conversely, the dynamic inflation target featured a time-varying slope but a constant target intercept.

Welfare gains and inflation bands. We now analyze the welfare gains achieved under our dynamic inflation target relative to another mechanism that many central banks use in practice, inflation caps (or bands). Concretely, we define a time-varying inflation cap as a mechanism where the central bank is subject to the constraint $\pi_t \leq \pi_t^*$ but faces no other constraints. It is immediate that under an inflation cap, the central bank's optimal policy is to adopt discretionary policy with $\pi_t = \theta_t + \frac{1}{\alpha\kappa}$ when this is interior, and to set inflation equal to the cap otherwise. Thus to first order, welfare gains are[29]

$$\mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t \alpha (\pi_t - \theta_t)^2 = \mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t \alpha \left[\frac{1}{(\alpha\kappa)^2} \mathbf{1}_{\theta_t + \frac{1}{\alpha\kappa} < \pi_t^*} + (\pi_t^* - \theta_t)^2 \mathbf{1}_{\theta_t + \frac{1}{\alpha\kappa} \geq \pi_t^*} \right].$$

Relative to a dynamic inflation target, a time-varying inflation cap mechanism results in welfare losses throughout the entire state space. In the region where the inflation cap does not bind, that is $\theta_t + \frac{1}{\alpha\kappa} < \pi_t^*$, the central bank's inflationary bias drives policy, leading to excess inflation and welfare losses. When the cap binds, that is $\theta_t + \frac{1}{\alpha\kappa} \geq \pi_t^*$, losses can result either from too little inflation or too much inflation. When $\theta_t < \pi_t^*$ but the cap binds, the central bank chooses too much inflation relative to the constrained efficient policy. On the other hand when $\theta_t > \pi_t^*$, constrained efficient inflation is actually above the cap, but the central bank is bound by the inflation cap. This leads to losses from too little inflation.

An inflation cap therefore implies welfare losses at everyone point in the state space except for the knife-edged point $\theta_t = \pi_t^*$ where the cap binds and is equal to the constrained efficient level of inflation. If the inflation cap is always updated one period in advance, then (from the first order approximation) the optimal inflation cap solves

$$\pi_t^* = \mathbb{E}_{t-1} \left[\theta_t \,\middle|\, \theta_t + \frac{1}{\alpha\kappa} > \pi_t^* \right],$$

which balances the losses from having too much inflation against having too little inflation is the region where the cap binds.

6 Extensions

In Section 4, we showed that a dynamic inflation target can implement the constrained efficient level of inflation. In this section, we study two main extensions: a conservative central banker, and costly enforcement.

6.1 Conservative Central Banker

Our first main extension is to allow for a conservative central banker as in Rogoff (1985). In particular, our conservative central banker places a greater penalty on inflation than the

[29] To obtain this, we take the first order approximation, $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\left(-\alpha(\pi_t - \theta_t) + \frac{1}{\kappa} \right) d\pi_t + \frac{-\beta}{\kappa} d\pi_{t+1} \right]$, and substitute in accordingly.

government. After appropriate intertemporal rearrangement of terms, we represent this by assuming central bank preferences equal to

$$V_t = U_t - c(\pi_t - \mathbb{E}_{t-1}\pi_t),$$

where as before U_t denotes the preferences of society and the government, and where c is the constant linear cost to the conservative central banker of inflation exceeding firm inflation expectations. We obtain the following minor modification of our main result.

Proposition 11. *Suppose that there is a conservative central banker, with flow utility $V_t = U_t - c(\pi_t - \mathbb{E}_{t-1}\pi_t)$. Then, the constrained efficient allocation can be implemented by a dynamic inflation target, with $b_{t-1} = \nu_{t-1} - c$ and $\tau_{t-1} = \mathbb{E}_{t-1}\pi_t$.*

Proposition 11 indicates that the presence of a conservative central banker with constant "conservativeness" c does not change the fundamental need for a dynamic inflation target when the flexibility ν_{t-1} of the target is changing over time. This is intuitive, since the inflation penalty applied by the conservative central banker is fixed over time. While a conservative central banker increases the flexibility of the target, in the sense that $b_{t-1} = \nu_{t-1} - c < \nu_{t-1}$, the total implied inflation penalty $b_{t-1} + c$, accounting for the conservative central banker's preferences, is ν_{t-1}, precisely as before. As a result, having a conservative central banker does not change the time variation in the optimal mechanism. Instead, it only changes the *average* extent to which inflation penalties are provided implicitly by the preferences of the conservative central banker, rather than explicitly by the dynamic inflation target.

Proposition 11 further suggests that an alternative implementation of the dynamic inflation target would be to have a time-varying central banker, whose preferences at date t were $c_t = \nu_{t-1}$. This implies that at date $t-1$, if $\nu_{t-1} > \nu_{t-2}$ then a more dovish central banker at date $t-1$ should be replaced by a more hawkish central banker at time t, and vice versa if $\nu_{t-1} < \nu_{t-2}$. Just as the dynamic inflation target is updated one period in advance, the appointment of a new central banker would also be announced one period in advance.

Importantly, just as a fixed central bank under the optimal mechanism was tasked with updating its own target, in an implementation with time varying conservativeness a central banker would be tasked with appointing her own replacement one period in advance (or at the least, would be responsible for naming her successor). However, this institutional arrangement is not typical (if used at all) in practice. For example, in the U.S.

the president is tasked with appointing members of the Board of Governors, who must then be confirmed by the Senate.

6.2 Costly Enforcement

Our second main extension is to allow for costly enforcement of the mechanism. We capture the social cost of implementing and enforcing a monetary policy mechanism by assuming that transfers are now costly to the government. Social preferences of the government are now

$$\max \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t\left(U_t(\pi_t,\mathbb{E}_t[\pi_{t+1}|\tilde{\theta}_t],\theta_t)-\kappa T_t\right)\right], \tag{20}$$

where $\kappa \geq 0$ captures the importance of this social cost.[30] In conjunction, we introduce the central bank participation constraint, given by

$$\mathcal{W}_0 \geq 0, \tag{21}$$

where without loss of generality we have normalized the outside option in the participation constraint to 0.[31] Recall that a mechanism is a mapping $(\pi_t, T_t) : \Theta^t \to \mathbb{R}^2$.

We leave detailed analysis to Appendix B, and directly characterize the result on the second best allocation. We solve for the optimal relaxed mechanism, which is subject to local incentive compatibility but not global incentive compatbility.

Proposition 12. *The solution to an optimal allocation rule of the relaxed problem is given by the first-order conditions*

$$\underbrace{\beta\frac{\partial U_t}{\partial \pi_t}+\frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}\pi_t}}_{\text{Constrained Efficient Trade-Off}} = \underbrace{\frac{\kappa}{1+\kappa}\left[\underbrace{\Gamma_{t-1}\frac{\partial^2 U_{t-1}}{\partial\theta_{t-1}\partial\mathbb{E}_{t-1}\pi_t}}_{\text{Persistence + Time Consistency}}+\underbrace{\beta\Gamma_t\frac{\partial^2 U_t}{\partial\theta_t\partial\pi_t}}_{\text{Persistence}}\right]}_{\text{Costly Transfers}} \forall t \geq 1$$

$$\frac{\partial U_0}{\partial \pi_0} = \frac{\kappa}{1+\kappa}\Gamma_0\frac{\partial^2 U_0}{\partial\theta_0\partial\pi_0}$$

[30] This corresponds to a standard (quasilinear) transferable utility model. As usual, T_t may also correspond to non-quasilinear utilities, provided they are transferable in this form.

[31] In Appendix B, we show that a dynamic inflation target is an optimal mechanism when there is instead an average participation constraint, $\mathbb{E}\mathcal{W}_0 \geq 0$.

where $\Gamma_t(\theta^t)$ is given by the recursive sequence

$$\Gamma_t(\theta^t) = \Gamma_{t-1}(\theta^{t-1}) \frac{1 - F(\theta_t|\theta_{t-1})}{f(\theta_t|\theta_{t-1})} \mathbb{E}_{t-1}\left[\frac{\partial f(s_t|\theta_{t-1})/\partial\theta_{t-1}}{f(s_t|\theta_{t-1})} \middle| s_t \geq \theta_t \right] \tag{22}$$

with initial condition $\Gamma_0(\theta^0) = \frac{1-F(\theta_0)}{f(\theta_0)}$.

If transfers were not costly ($\kappa = 0$), the optimal allocation would be the constrained efficient allocation, and the dynamic inflation target would be an optimal mechanism. However, there are two new effects arising from the fact that transfers are now costly. These effects reflect both the persistence of information and the time consistency problem. Because the central bank earns an information rent, implementing a given allocation rule is now costly. Suppose that on the margin, the government changes the allocation π_t, holding all else fixed. This affects the required transfer at two points. First, it affects the required transfer in period t. Information rents at date t are related to $\frac{\partial U_t}{\partial \theta_t}$, so that the effect of the allocation on information rents is related to the cross partial $\frac{\partial^2 U_t}{\partial\theta_t \partial\pi_t}$. Assuming that $\frac{\partial^2 U_t}{\partial\theta_t \partial\pi_t} > 0$ and that $\Gamma_t > 0$, then an increase in π_t increases the information rent $\frac{\partial U_t}{\partial \theta_t} > 0$, which in turn increases the required transfer, generating a cost for the government. This is a standard effect that exists even without a time consistency problem.

However, changes in π_t also affect the required transfers at date $t-1$, due to the Phillips curve relationship. In particular, the information rent $\frac{\partial U_{t-1}}{\partial \theta_{t-1}}$ now depends on π_t through inflation expectations. In particular if $\Gamma_{t-1} > 0$ and $\frac{\partial^2 U_{t-1}}{\partial\theta_{t-1} \mathbb{E}_{t-1}\pi_t}$, then the effect is analogous. An increase in π_t causes an increase in $\frac{\partial U_{t-1}}{\partial \theta_{t-1}}$, which increases the required transfer. In other words, this second effect is an interaction between the time consistency problem and the information persistence problem.

In either case, note that when the RHS is positive, we have $\frac{\partial U_t}{\partial \pi_t} + \frac{1}{\beta}\frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}\pi_t} > 0$. This implies that from the social perspective (accounting for time consistency), it would be desirable to increase π_t further to reach the constrained efficient allocation. However, it is costly to provide the correct incentive scheme to do so, and as a result the time consistency problem is not fully corrected. Costly enforcement in this event reduces the ability to combat the time consistency problem.

Lastly, the initial condition $\Gamma_0(\theta^0) = \frac{1-F(\theta_0)}{f(\theta_0)}$ implies a conventional first-order condition for period-0 inflation

$$\frac{\partial U_0}{\partial \pi_0} - \frac{\kappa}{1+\kappa}\Gamma_0(\theta^0)\frac{\partial^2 U_0}{\partial\theta_0 \partial\pi_0} = 0$$

This reflects a standard virtual value relationship. In our model, the component $(1 - F(\theta_0))/f(\theta_0) > 0$ of virtual value is relevant at both date 0 and date 1, due to the time consistency problem. Moreover, this value $(1 - F(\theta_0))/f(\theta_0) > 0$ is persistent over time through the evolution of Γ_t. In other words, if it is large initially, then Γ_t is also larger in magnitude along every history. Notice that absent persistent shocks, $\Gamma_t = 0$ for all $t \geq 2$ along any history. In other words absent persistent shocks, all rents are extracted by distorting allocations at dates 0 and 1, but thereafter the allocation reverts to constrained efficiency and hence to implementation via a dynamic inflation target.

It should be noted that the constrained efficient allocation is still *implementable*, but is no longer optimal. In particular, at the constrained efficient allocation, the informational effects still exactly offset one another, and a dynamic inflation target still implements that allocation. However, at the constrained efficient allocation the marginal benefit of higher inflation is zero, while the marginal cost of the enforcement mechanism is not. As a result, the government no longer finds it optimal to implement the constrained efficient allocation that would have arisen under zero costs. This helps to understand why the optimal mechanism here deviates from a precise dynamic inflation target: once moved away from the constrained efficient allocation, the informational effect on firms no longer exactly offsets the informational effect on government transfers.

Nevertheless, the properties of the optimum still bear similarities to the dynamic inflation target. In particular, the time consistency term is still a date $t-1$ adapted constant given the Phillips curve relationship, generating a component of the optimal mechanism that resembles the dynamic target. Moreover, notice that even under Proposition 12, there are a set of sufficient statistics for the shock history θ^{t-1}, which are now the tuple $(\nu_{t-1}, \Gamma_{t-1}, \theta_{t-1}, \frac{\partial^2 U_{t-1}}{\partial \theta_{t-1} \partial \mathbb{E}_{t-1} \pi_t})$ of date $t-1$ adapted constants. These sufficient statistics can be used to derive the allocation and transfer rule at date t.

Finally, we can use Proposition 12 to show that the optimal mechanism in fact reverts to a dynamic inflation target at both extremes of the shock distribution. In other words, there is both a no top distortion *and* a no bottom distortion result: if at period t we have $\theta_t = \overline{\theta}$ or $\theta_t = \underline{\theta}$, then $\Gamma_t = 0$, and hence $\Gamma_{t+s} = 0 \; \forall t \geq s$. This implies that the entire costly transfer term disappears, at which point the optimal mechanism reverts to the dynamic inflation target.

Corollary 13. If $\theta_t \in \{\underline{\theta}, \overline{\theta}\}$, then $\Gamma_{t+s} = 0 \; \forall s \geq 0$ and the optimal allocation at dates $t+s$ ($s \geq 1$) can be implemented by a dynamic inflation target.

The no top distortion result closely resembles normal top distortion results in the absence of time consistency. In our model, the time consistency problem implies that the optimal allocation rule that is the constrained efficient allocation rule, rather than the rule under discretion. Moreover due to information persistence, we also have a no distortion at the bottom result.[32] As a result, not only does the optimal mechanism has a component that resembles the dynamic inflation target throughout the shock distribution, but it reverts fully to the dynamic inflation target at the limits of the distribution.

7 Conclusion

We provide a theory of how a central bank should update its inflation target in the presence of persistent economic shocks. We show that a dynamic inflation targeting mechanism can implement the constrained efficient level of inflation, and that when enforcement costs of using the mechanism are negligible it is an optimal mechanism. Both the level and flexibility of the target may be updated. A key property is that the dynamic inflation target is updated *one period in advance,* which mitigates the time consistency problem in target setting. This suggests that a key property of target efficacy is advanced commitment. The results provide guidance on the mechanism by which to adjust an inflation target without exacerbating the underlying time consistency problem, suggesting that a mechanism of adjustment at restricted points in time, for example every five years as with the Bank of Canada, would be the desirable adjustment method.

[32] See Pavan et al. (2014) for related results.

References

Amador, M., I. Werning, and G. M. Angeletos (2006). Commitment vs. flexibility. *Econometrica 74*(2), 365–396.

Andrade, P., J. Gal\'\i, H. Le Bihan, and J. Matheron (2018). The optimal inflation target and the natural rate of interest. Technical report, National Bureau of Economic Research.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006). Signaling in a global game: Coordination and policy traps. *Journal of Political economy 114*(3), 452–484.

Athey, S., A. Atkeson, and P. J. Kehoe (2005). The optimal degree of discretion in monetary policy. *Econometrica 73*(5), 1431–1475.

Backus, D. and J. Driffill (1985). Inflation and reputation. *The American Economic Review 75*(3), 530–538.

Ball, L. M. (2013). The case for four percent inflation. Technical report, working paper.

Barro, R. J. and D. B. Gordon (1983, jan). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics 12*(1), 101–121.

Beshears, J., J. Choi, C. Clayton, C. Harris, D. Laibson, and B. Madrian (2020). Optimal Illiquidity.

Blanchard, O. (2016). The Phillips Curve: Back to the'60s? *American Economic Review 106*(5), 31–34.

Blanchard, O., G. Dell'Ariccia, and P. Mauro (2010). Rethinking macroeconomic policy. *Journal of Money, Credit and Banking 42*, 199–215.

Brainard, L. (2015). No Title. Technical report, Board of Governors of the Federal Reserve System. Speech delivered at the 57th National Association for Business Economics Annual Meeting.

Campbell, J. R., C. L. Evans, J. D. M. Fisher, A. Justiniano, C. W. Calomiris, and M. Woodford (2012). Macroeconomic effects of federal reserve forward guidance [with comments and discussion]. *Brookings papers on economic activity*, 1–80.

Canzoneri, M. B. (1985). Monetary policy games and the role of private information. *American Economic Review 75*(5), 1056–1070.

Clarida, R. H. (2019). The Federal Reserve's Review of Its Monetary Policy Strategy, Tools, and Communication Practices. Technical report, Board of Governors of the Federal Reserve System. Speech delivered at the 2019 U.S. Monetary Policy Forum, sponsored by the Initiative on Global Markets at the University of Chicago Booth School of Business, held in New York, February 22.

Coibion, O., Y. Gorodnichenko, and J. Wieland (2012). The optimal inflation rate in New Keynesian models: should central banks raise their inflation targets in light of the zero lower bound? *Review of Economic Studies 79*(4), 1371–1406.

Cukierman, A. and A. H. Meltzer (1986). A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica: journal of the econometric society*, 1099–1128.

Del Negro, M., M. Lenza, G. E. Primiceri, and A. Tambalotti (2020). What's up with the Phillips Curve? Technical report, National Bureau of Economic Research.

DellaVigna, S. and U. Malmendier (2004). Contract Design and Self-Control: Theory and Evidence. *Quarterly Journal of Economics 119*(2), 353–402.

Eggertsson, G. B., N. R. Mehrotra, and J. A. Robbins (2019). A model of secular stagnation: Theory and quantitative evaluation. *American Economic Journal: Macroeconomics 11*(1), 1–48.

Farhi, E. and I. Werning (2013). Insurance and taxation over the life cycle. *Review of Economic Studies 80*(2), 596–635.

Friedman, M. (1972). Have monetary policies failed? *The American Economic Review 62*(1/2), 11–18.

Gal\'\i, J. and L. Gambetti (2019). Has the US wage Phillips curve flattened? A semi-structural exploration. Technical report, National Bureau of Economic Research.

Galperti, S. (2015). Commitment, Flexibility, and Optimal Screening of Time Inconsistency. *Econometrica 83*(4), 1425–1465.

Goodhart, C. A. E. (2001). Monetary transmission lags and the formulation of the policy decision on interest rates. In *Challenges for Central Banking*, pp. 205–228. Springer.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American economic review 95*(1), 425–436.

Halac, M. and P. Yared (2014). Fiscal Rules and Discretion Under Persistent Shocks. *Econometrica 82*(5), 1557–1614.

Halac, M. and P. Yared (2018). Fiscal Rules and Discretion in a World Economy. *American Economic Review 108*(8), 2305–2334.

Halac, M. and P. Yared (2019). Instrument-Based vs. Target-Based Rules.

Havranek, T. and M. Rusnak (2012). Transmission lags of monetary policy: A meta-analysis.

Kiley, M. T. and J. M. Roberts (2017). Monetary policy in a low interest rate world. *Brookings Papers on Economic Activity 2017*(1), 317–396.

Krishnamurthy, A. and A. Vissing-Jorgensen (2012). The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy. *Brookings Papers on Economic Activity Fall 2011*, 215.

Krugman, P. (2014). Inflation targets reconsidered. In *Draft paper for ECB Sintra conference*.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of monetary economics 47*(3), 523–544.

Kydland, F. E. and E. C. Prescott (1977, jun). Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy 85*(3), 473–491.

Laubach, T. and J. C. Williams (2016). Measuring the natural rate of interest redux. *Business Economics 51*(2), 57–67.

L'Huillier, J.-P. and R. Schoenle (2019). Raising the Target: How Much Extra Room Does It Really Give?

Lucca, D. O. and E. Moench (2015). The pre-FOMC announcement drift. *The Journal of finance 70*(1), 329–371.

McDermott, J. and R. Williams (2018). Inflation Targeting in New Zealand: an experience in evolution. Technical report.

Pavan, A., I. Segal, and J. Toikka (2014). Dynamic Mechanism Design: A Myersonian Approach. *Econometrica 82*(2), 601–653.

Persson, T. and G. Tabellini (1993, dec). Designing institutions for monetary stability. *Carnegie-Rochester Confer. Series on Public Policy 39*(C), 53–84.

Powell, J. H. (2019). Monetary policy in the post-crisis era. Speech at "Bretton Woods: 75 Years Later - Thinking about the Next 75".

Powell, J. H. (2020). New Economic Challenges and the Fed's Monetary Policy Review. Technical report, Board of Governors of the Federal Reserve System. Speech delivered at "Navigating the Decade Ahead: Implications for Monetary Policy," an economic policy symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.

Rogoff, K. (1985). The Optimal Degree of Commitment to an Intermediate Monetary Target. *The Quarterly Journal of Economics 100*(4), 1169–1189.

Romer, C. D. and D. H. Romer (2000, jun). Federal Reserve Information and the Behavior of Interest Rates. *American Economic Review 90*(3), 429–457.

Rubbo, E. Networks, Phillips Curves and Monetary Policy.

Rudebusch, G. and L. E. O. Svensson (1999). Policy rules for inflation targeting. In *Monetary policy rules*, pp. 203–262. University of Chicago Press.

Sargent, T. J. and N. Wallace (1975). " Rational" expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of political economy 83*(2), 241–254.

Schmitt-Grohé, S. and M. Uribe (2010). The optimal rate of inflation. In *Handbook of monetary economics*, Volume 3, pp. 653–722. Elsevier.

Sleet, C. (2001). On credible monetary policy and private government information. *Journal of Economic Theory 99*(1-2), 338–376.

Svensson, L. E. (2010). Inflation targeting. *Handbook of Monetary Economics 3*(C), 1237–1302.

Svensson, L. E. O. (1995). Optimal inflation targets,conservative'central banks, and linear inflation contracts. Technical report, National Bureau of Economic Research.

Taylor, J. B. and V. Wieland (2012). Surprising comparative properties of monetary models: results from a new model database. *Review of Economics and Statistics 94*(3), 800–816.

Waki, Y., R. Dennis, and I. Fujiwara (2018). The optimal degree of monetary discretion in a new Keynesian model with private information. *Theoretical Economics 13*(3), 1319–1367.

Walsh, C. E. (1995). Optimal contracts for central bankers. *American Economic Review 85*(1), 150–167.

A Proofs

A.1 Proof of Proposition 1

Under full information, we can apply law of iterated expectations ($\mathbb{E}\mathbb{E}_{t-1}\pi_t = \mathbb{E}\pi_t$) and rearrange equation (3) to

$$\mathbb{E}\frac{1}{\kappa}\left[\frac{1}{\theta_0}\pi_0 + \sum_{t=1}^{\infty}\beta^t\left[\frac{1}{\theta_t} - \frac{1}{\theta_{t-1}}\right]\pi_t\right].$$

That $\pi_0 = \overline{\pi}$ follows immediately. From here, the derivative of the objective in inflation $\pi_t(\theta^t)$ following history θ^t for $t \geq 1$ is

$$\frac{1}{\kappa}\beta^t\left[\frac{1}{\theta_t} - \frac{1}{\theta_{t-1}}\right]f(\theta^t).$$

If $\theta_{t-1} = \underline{\theta}$, then this is weakly negative for any $\theta_t \in \{\underline{\theta}, \overline{\theta}\}$, and so an optimal policy is the corner solution $\pi_t = \underline{\pi}$ for any θ_t. Likewise if $\theta_{t-1} = \overline{\theta}$, then this is weakly positive, and so an optimal policy is the corner solution $\pi_t = \overline{\pi}$ for any θ_t.

A.2 Proof of Proposition 2

Suppose that all future central banks adopt the allocation $\pi_{t+s} = \overline{\pi}$ for $s \geq 1$. Then, $\mathbb{E}_t\pi_{t+1} = \overline{\pi}$ for all t and θ_t, and hence the central bank objective at date t is optimization equivalent to $\frac{1}{\kappa}\frac{1}{\theta_t}\pi_t$. Hence, $\pi_t = \overline{\pi}$ is an optimal choice.

A.3 Proof of Proposition 3

Consider the central bank at date t, whose welfare function given an inherited target π_t^* is

$$\mathcal{W}_t = \frac{1}{\kappa}\left[\frac{1}{\theta_t}\pi_t^* + \sup_{\pi_{t+1}^*}\left\{-\beta\frac{1}{\theta_t}\pi_{t+1}^* + \beta\mathbb{E}_t\mathcal{W}_{t+1}\right\}\right],$$

given that firm expectations must be $\mathbb{E}_t\pi_{t+1} = \pi_{t+1}^*$. From here, suppose all future central banks are choosing π_{t+s+1}^* according to the conjectured policy for $s \geq 1$. Then, we have

$$\arg\sup_{\pi_{t+1}^*}\left\{-\beta\frac{1}{\theta_t}\pi_{t+1}^* + \beta\mathbb{E}_t\mathcal{W}_{t+1}\right\} = \arg\sup_{\pi_{t+1}^*}\left\{-\beta\frac{1}{\theta_t}\pi_{t+1}^* + \beta\mathbb{E}_t\frac{1}{\theta_{t+1}}\pi_{t+1}^*\right\},$$

giving the result.

A.4 Proof of Proposition 4

Suppose that all future central banks update targets at date $t+s$ to $\pi^*_{t+s} = \overline{\pi}$ for $s \geq 1$. Then, $\mathbb{E}_t \pi_{t+1} = \overline{\pi}$, and hence $\pi^*_t = \pi_t = \overline{\pi}$ is an optimal choice, as in the proof of Proposition 2.

A.5 Proof of Proposition 5

Under full information, the objective function of the government is

$$\sup_{\pi_t} E_0 \sum_{t=0}^{\infty} \beta^t U_t \left(\pi_t, E_t \left[\pi_{t+1} | \theta_t\right], \theta_t\right),$$

from which the first order conditions follow immediately.

A.6 Proof of Proposition 6

Consider an equilibrium where the central bank sets π^*_t as conjectured period by period. Then, future inflation policies and expectations are taken as given, and the central bank's objective at date t collapses to

$$\sup_{\pi^*_t} U_t \left(\pi^*_t, E_t \left[\pi_{t+1} | \theta_t\right], \theta_t\right),$$

from which the result follows.

A.7 Proof of Proposition 7

The proof strategy is as follows. First, we derive the relevant envelope condition associated with local incentive compatibility, which defines necessary conditions on the value function associated with an incentive compatible mechanism.[33] We then show that the value function generated by our proposed mechanism satisfies this envelope condition.

Envelope Condition. Suppose that the central bank has a history $\tilde{\theta}^{t-1}$ of reports and a history θ^t of true types at date t. Given a mechanism with transfer rule $T_t(\tilde{\theta}_t)$ and allocation rule $\pi_t(\tilde{\theta}_t)$, the value function of a central bank that has truthfully reported in the past,

[33]This portion of the argument follows from the arguments in Farhi and Werning (2013), or more generally from Pavan et al. (2014), but we state it out for completeness and for clarity.

assuming truthful reporting in the future, is given by

$$\mathcal{W}_t(\theta^t) = \max_{\tilde{\theta}_t} T_t + U_t\left(\pi_t, \mathbb{E}_t\left[\pi_{t+1}|\tilde{\theta}_t\right], \theta_t\right) + \beta \mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^t, \tilde{\theta}_t, \theta_{t+1})\middle|\theta_t\right]$$

Notice that the Phillips Curve expectation $\mathbb{E}_t\left[\pi_{t+1}|\tilde{\theta}_t\right]$ is based on the date t reported type, not the date t true type. Furthermore, notice that $\mathcal{W}_{t+1}$ depends on the reported type $\tilde{\theta}_t$, but not on the true type θ_t. This is because flow utility at dates $t+s$ ($s \geq 0$) do not depend on past true types and because the shock structure is Markov. This implies that we can in fact write $\mathcal{W}_{t+1}(\theta^{t-1}, \tilde{\theta}_t, \theta_{t+1})$. As a result, the Envelope Condition in the true type θ_t, evaluated at truthful reporting $\tilde{\theta}_t = \theta_t$, is

$$\frac{\partial \mathcal{W}_t(\theta^t)}{\partial \theta_t} = \frac{\partial U_t\left(\pi_t, \mathbb{E}_t\left[\pi_{t+1}|\tilde{\theta}_t\right], \theta_t\right)}{\partial \theta_t} + \beta \frac{\partial \mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t-1}, \tilde{\theta}_t, \theta_{t+1})\middle|\theta_t\right]}{\partial \theta_t}$$

where we have

$$\begin{aligned}\frac{\partial \mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t-1}, \tilde{\theta}_t, \theta_{t+1})\middle|\theta_t\right]}{\partial \theta_t} &= \frac{\partial}{\partial \theta_t}\int_{\underline{\theta}}^{\overline{\theta}} \mathcal{W}_{t+1}(\theta^{t-1}, \tilde{\theta}_t, \theta_{t+1}) f(\theta_{t+1}|\theta_t) d\theta_{t+1} \\ &= \mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t-1}, \tilde{\theta}_t, \theta_{t+1}) \frac{\partial f(\theta_{t+1}|\theta_t)/\partial \theta_t}{f(\theta_{t+1}|\theta_t)}\middle|\theta_t\right]\end{aligned}$$

Substituting in and evaluating at truthful reporting, we obtain

$$\frac{\partial \mathcal{W}_t(\theta^t)}{\partial \theta_t} = \frac{\partial U_t\left(\pi_t, \mathbb{E}_t\left[\pi_{t+1}|\theta_t\right], \theta_t\right)}{\partial \theta_t} + \beta \mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t+1}) \frac{\partial f(\theta_{t+1}|\theta_t)/\partial \theta_t}{f(\theta_{t+1}|\theta_t)}\middle|\theta_t\right]$$

which provides a conventional envelope condition for incentive compatibility. For clarity, note that $\frac{\partial U_t(\pi_t, \mathbb{E}_t[\pi_{t+1}|\theta_t], \theta_t)}{\partial \theta_t}$ is the derivative of U_t in the direct type θ_t, but *not* including the Phillips Curve expectation, which is the derivative in the reported type.

Verifying the Envelope Condition. We now verify the value function under our mechanism satisfies the envelope condition. Our mechanism has a transfer rule $T_t(\theta^t) = -\nu_{t-1}(\theta^{t-1})\Big(\pi_t(\theta^t) - \mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]\Big)$ and an allocation rule given by the constrained efficient allocation of Proposition 5. The value function associated with this mechanism is

$$\mathcal{W}_t(\theta^t) = -\nu_{t-1}\bigg(\pi_t - \mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]\bigg) + U_t\left(\pi_t, \mathbb{E}_t\left[\pi_{t+1}|\theta_t\right], \theta_t\right) + \beta\mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t+1})\middle|\theta_t\right]$$

where $\nu_{t-1}, \pi_t, \mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]$) are the constrained efficient values associated with Proposition 5, given the realized shock history. From here, recall that ν_{t-1} and $\mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]$ are only functions of θ^{t-1}. Therefore, $\frac{\partial\nu_{t-1}}{\partial\theta_t} = \frac{\partial\mathbb{E}_{t-1}[\pi_t|\theta_{t-1}]}{\partial\theta_t} = 0$. Thus differentiating the value function in θ_t, we have

$$\begin{aligned}\frac{\partial\mathcal{W}_t(\theta^t)}{\partial\theta_t} =& \frac{\partial U_t}{\partial\theta_t} + \beta\mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t+1})\frac{\partial f(\theta_{t+1}|\theta_t)/\partial\theta_t}{f(\theta_{t+1}|\theta_t)}\middle|\theta_t\right] \\ & - \nu_{t-1}\frac{\partial\pi_t}{\partial\theta_t} + \frac{\partial U_t}{\partial\pi_t}\frac{\partial\pi_t}{\partial\theta_t} + \frac{\partial U_t}{\partial\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}\frac{d\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}{d\theta_t} + \beta\mathbb{E}_t\left[\frac{\partial\mathcal{W}_{t+1}(\theta^{t+1})}{\partial\theta_t}\middle|\theta_t\right]\end{aligned}$$

The first line on the RHS are the terms associated with the envelope condition. The second line are derivatives that arise because in equilibrium, the reported type equals the true type, and we have evaluated the value function given truthful reporting. It therefore remains to show that the second line sums to zero and hence our mechanism satisfies the required envelope condition.

It is helpful to write out the continuation value function $\mathcal{W}_{t+1}$ in sequence notation. Iterating forward, we obtain

$$\begin{aligned}\mathcal{W}_{t+1}(\theta^{t+1}) =& -\nu_t\bigg(\pi_{t+1} - \mathbb{E}_t[\pi_{t+1}|\theta_t]\bigg) \\ & - \mathbb{E}_{t+1}\left[\sum_{s=1}^{\infty}\beta^s\nu_{t+s}\bigg(\pi_{t+1+s} - \mathbb{E}_{t+s}[\pi_{t+1+s}|\theta_{t+s}]\bigg)\middle|\theta_{t+1}\right] \\ & + \mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^s U_{t+1+s}\left(\pi_{t+1+s}, \mathbb{E}_{t+1+s}\left[\pi_{t+2+s}|\theta_{t+1+s}\right], \theta_{t+1+s}\right)\middle|\theta_{t+1}\right]\end{aligned}$$

The first two lines on the RHS are total expected discounted value arising from transfers. The third line on the RHS is total expected discounted value arising from flow utility.

Notice from here that the second line is equal to zero. To see this, applying Law of Iterated Expectations, when $s \geq 1$ we have

$$\mathbb{E}_{t+1}\left[\nu_{t+s}\pi_{t+1+s}|\theta_{t+1}\right] = \mathbb{E}_{t+1}\left[\mathbb{E}_{t+s}\left[\nu_{t+s}\pi_{t+1+s}\middle|\theta_{t+s}\right]|\theta_{t+1}\right] = \mathbb{E}_{t+1}\left[\nu_{t+s}\mathbb{E}_{t+s}\left[\pi_{t+1+s}\middle|\theta_{t+s}\right]|\theta_{t+1}\right]$$

since ν_{t+s} is a function only of θ^{t+s}, and so is known at date $t+s$. As a result, the second line is zero, and we can write

$$\mathcal{W}_{t+1}(\theta^{t+1}) = -\nu_t\Big(\pi_{t+1} - \mathbb{E}_t[\pi_{t+1}|\theta_t]\Big) \\ + \mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^s U_{t+1+s}\left(\pi_{t+1+s}, \mathbb{E}_{t+1+s}\left[\pi_{t+2+s}|\theta_{t+1+s}\right], \theta_{t+1+s}\right)\Big|\theta_{t+1}\right]$$

From here, we differentiate the continuation value $\mathcal{W}_{t+1}(\theta^{t+1})$ in the date t type θ_t, yielding

$$\frac{\partial \mathcal{W}_{t+1}(\theta^{t+1})}{\partial \theta_t} = -\frac{\partial \nu_t}{\partial \theta_t}\Big(\pi_{t+1} - \mathbb{E}_t[\pi_{t+1}|\theta_t]\Big) - \nu_t\left(\frac{\partial \pi_{t+1}}{\partial \theta_t} - \frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}\right) \\ + \mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^s\left(\frac{\partial U_{t+1+s}}{\partial \pi_{t+1+s}}\frac{\partial \pi_{t+1+s}}{\partial \theta_t} + \frac{\partial U_{t+1+s}}{\partial \mathbb{E}_{t+1+s}\left[\pi_{t+2+s}|\theta_{t+1+s}\right]}\mathbb{E}_{t+1+s}\left[\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\right)\right.$$

Notice in the above derivation that only the first line includes a total derivative of firm expectations, $\frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}$, which accounts for the changes in probability density. All later lines only include the direct change in inflation policy. This is because conditional expectations at date $t+1$ are taken with respect to θ_{t+1}, not θ_t.

We now rearrange the first term on the second line as follows. In particular, we write

$$\sum_{s=0}^{\infty}\beta^s\frac{\partial U_{t+1+s}}{\partial \pi_{t+1+s}}\frac{\partial \pi_{t+1+s}}{\partial \theta_t} = \frac{\partial U_{t+1}}{\partial \pi_{t+1}}\frac{\partial \pi_{t+1}}{\partial \theta_t} + \sum_{s=0}^{\infty}\beta^{s+1}\frac{\partial U_{t+2+s}}{\partial \pi_{t+2+s}}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}$$

which extracts the first element of the sum, and relabels the remainder of the sum to continue to start from $s=0$. Substituting back in, we obtain

$$\frac{\partial \mathcal{W}_{t+1}(\theta^{t+1})}{\partial \theta_t} = -\frac{\partial \nu_t}{\partial \theta_t}\Big(\pi_{t+1} - \mathbb{E}_t[\pi_{t+1}|\theta_t]\Big) - \nu_t\left(\frac{\partial \pi_{t+1}}{\partial \theta_t} - \frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}\right) + \frac{\partial U_{t+1}}{\partial \pi_{t+1}}\frac{\partial \pi_{t+1}}{\partial \theta_t} \\ + \mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(\frac{\partial U_{t+2+s}}{\partial \pi_{t+2+s}}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}\right.\right. \\ \left.\left. + \frac{1}{\beta}\frac{\partial U_{t+1+s}}{\partial \mathbb{E}_{t+1+s}\left[\pi_{t+2+s}|\theta_{t+1+s}\right]}\mathbb{E}_{t+1+s}\left[\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\right)\Big|\theta_{t+1}\right].$$

By definition, we have $\nu_{t+s+1} = -\frac{1}{\beta}\frac{\partial U_{t+1+s}}{\partial \mathbb{E}_{t+1+s}[\pi_{t+2+s}|\theta_{t+1+s}]}$, given the allocation rule we are using in constructing the value function is the constrained efficient allocation rule. By

Proposition 5, we also have $\frac{\partial U_{t+2+s}}{\partial \pi_{t+2+s}} = \nu_{t+s+1}$ for the same reason. Therefore, we can write

$$\mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(\frac{\partial U_{t+2+s}}{\partial \pi_{t+2+s}}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}+\frac{1}{\beta}\frac{\partial U_{t+1+s}}{\partial \mathbb{E}_{t+1+s}\left[\pi_{t+2+s}|\theta_{t+1+s}\right]}\mathbb{E}_{t+1+s}\left[\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\right)\middle|\theta_{t+1}\right]$$
$$=\mathbb{E}_{t+1}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(\nu_{t+1+s}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}-\nu_{t+1+s}\mathbb{E}_{t+1+s}\left[\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\right)\middle|\theta_{t+1}\right]$$
$$=0$$

where the last line follows by Law of Iterated expectations,

$$\mathbb{E}_{t+1}\left[\nu_{t+1+s}\mathbb{E}_{t+1+s}\left[\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\middle|\theta_{t+1}\right]=\mathbb{E}_{t+1}\left[\mathbb{E}_{t+1+s}\left[\nu_{t+1+s}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}|\theta_{t+1+s}\right]\middle|\theta_{t+1}\right]$$
$$=\mathbb{E}_{t+1}\left[\nu_{t+1+s}\frac{\partial \pi_{t+2+s}}{\partial \theta_t}\middle|\theta_{t+1}\right].$$

Therefore, we obtain

$$\frac{\partial \mathcal{W}_{t+1}(\theta^{t+1})}{\partial \theta_t}=-\frac{\partial \nu_t}{\partial \theta_t}\left(\pi_{t+1}-\mathbb{E}_t[\pi_{t+1}|\theta_t]\right)-\nu_t\left(\frac{\partial \pi_{t+1}}{\partial \theta_t}-\frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}\right)+\frac{\partial U_{t+1}}{\partial \pi_{t+1}}\frac{\partial \pi_{t+1}}{\partial \theta_t}.$$

Finally, notice that as before, by Proposition 5 we have $\nu_t = \frac{\partial U_{t+1}}{\partial \pi_{t+1}}$, and therefore we can write

$$\frac{\partial \mathcal{W}_{t+1}(\theta^{t+1})}{\partial \theta_t}=-\frac{\partial \nu_t}{\partial \theta_t}\left(\pi_{t+1}-\mathbb{E}_t[\pi_{t+1}|\theta_t]\right)+\nu_t\frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}.$$

We are now ready to substitute back in to the expression for $\frac{\partial \mathcal{W}_t}{\partial \theta_t}$. Substituting back in, we have

$$\begin{aligned}\frac{\partial \mathcal{W}_t(\theta^t)}{\partial \theta_t}=&\frac{\partial U_t}{\partial \theta_t}+\beta\mathbb{E}_t\left[\mathcal{W}_{t+1}(\theta^{t+1})\frac{\partial f(\theta_{t+1}|\theta_t)/\partial\theta_t}{f(\theta_{t+1}|\theta_t)}\middle|\theta_t\right]\\&-\nu_{t-1}\frac{\partial \pi_t}{\partial \theta_t}+\frac{\partial U_t}{\partial \pi_t}\frac{\partial \pi_t}{\partial \theta_t}+\frac{\partial U_t}{\partial \mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}\frac{d\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}{d\theta_t}\\&+\beta\mathbb{E}_t\left[-\frac{\partial \nu_t}{\partial \theta_t}\left(\pi_{t+1}-\mathbb{E}_t[\pi_{t+1}|\theta_t]\right)+\nu_t\frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t}\middle|\theta_t\right]\end{aligned}$$

The arguments now are familiar. The first term on the third line is zero, since

$$\mathbb{E}_t\left[-\frac{\partial \nu_t}{\partial \theta_t}\left(\pi_{t+1}-\mathbb{E}_t[\pi_{t+1}|\theta_t]\right)\middle|\theta_t\right]=-\frac{\partial \nu_t}{\partial \theta_t}\mathbb{E}_t\left[\pi_{t+1}-\mathbb{E}_t[\pi_{t+1}|\theta_t]\middle|\theta_t\right]=0.$$

From here, we can rearrange terms to get

$$\frac{\partial \mathcal{W}_t(\theta^t)}{\partial \theta_t} = \frac{\partial U_t}{\partial \theta_t} + \beta \mathbb{E}_t \left[\mathcal{W}_{t+1}(\theta^{t+1}) \frac{\partial f(\theta_{t+1}|\theta_t)/\partial \theta_t}{f(\theta_{t+1}|\theta_t)} \middle| \theta_t \right]$$
$$+ \left[-\nu_{t-1} + \frac{\partial U_t}{\partial \pi_t} \right] \frac{\partial \pi_t}{\partial \theta_t} + \frac{\partial U_t}{\partial \mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]} \frac{d\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}{d\theta_t} + \beta \mathbb{E}_t \left[\nu_t \frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t} \middle| \theta_t \right]$$

By Proposition 5, we have $-\nu_{t-1} + \frac{\partial U_t}{\partial \pi_t} = 0$.[34] Likewise from the definition of ν_t, we have $\frac{\partial U_t}{\partial \mathbb{E}_t[\pi_{t+1}|\theta_t]} = -\beta \nu_t$. Therefore, we also have

$$\frac{\partial U_t}{\partial \mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]} \frac{d\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}{d\theta_t} + \beta \mathbb{E}_t \left[\nu_t \frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t} \middle| \theta_t \right] = -\beta \nu_t \frac{d\mathbb{E}_t\left[\pi_{t+1}|\theta_t\right]}{d\theta_t} + \beta \nu_t \frac{d\mathbb{E}_t[\pi_{t+1}|\theta_t]}{d\theta_t} = 0.$$

Thus, the entire second line is zero, and we are left with

$$\frac{\partial \mathcal{W}_t(\theta^t)}{\partial \theta_t} = \frac{\partial U_t}{\partial \theta_t} + \beta \mathbb{E}_t \left[\mathcal{W}_{t+1}(\theta^{t+1}) \frac{\partial f(\theta_{t+1}|\theta_t)/\partial \theta_t}{f(\theta_{t+1}|\theta_t)} \middle| \theta_t \right]$$

which is the required envelope condition. This concludes the proof.

A.8 Proof of Proposition 8

To first order, the welfare gains of an inflation perturbation from the static target is

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\partial U_t}{\partial \pi_t} d\pi_t + \frac{\partial U_t}{\partial \mathbb{E}_t \pi_{t+1}} d\pi_{t+1} \right].$$

From here, the first order condition of the central bank is $\nu^* = \frac{\partial U_t}{\partial \pi_t}$, while by definition $\frac{\partial U_t}{\partial \mathbb{E}_t \pi_{t+1}} = -\beta \nu_t^*$. We have $\frac{\partial U_0}{\partial \pi_0} = 0$, so that we have

$$\mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t \left[\nu^* - \nu_{t-1}^* \right] d\pi_t.$$

Finally, we have $\mathbb{E}_{t-1} d\pi_t = \tau_{t-1} - \mathbb{E}_{t-1} \pi_t^*$, giving the result.

[34]For completeness, note that when considering the date 0 value function, we have $\nu_{-1} = 0$ and have $\frac{\partial U_t}{\partial \pi_t} = 0$ by Proposition 5.

A.9 Proof of Proposition 9

Internalizing the NKPC into the loss function, we have

$$U_t = -\frac{1}{2}\pi_t^2 + \theta_t \frac{\pi_t - \beta^* \mathbb{E}_t \pi_{t+1}}{\kappa}.$$

From here, we have

$$\frac{\partial U_t}{\partial \pi_t} = -\pi_t + \frac{1}{\kappa}\theta_t$$
$$\frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}\pi_t} = -\frac{\beta^*}{\kappa}\theta_{t-1}$$

from which the results follow using Propositions 5 and 7.

A.10 Proof of Proposition 10

Internalizing the NKPC into the loss function, we have

$$U_t = -\frac{1}{2}\alpha(\pi_t - \theta_t)^2 + \frac{\pi_t - \beta \mathbb{E}_t \pi_{t+1}}{\kappa}$$

so that we have

$$\frac{\partial U_t}{\partial \pi_t} = -\alpha(\pi_t - \theta_t) + \frac{1}{\kappa}$$
$$\frac{\partial U_{t-1}}{\partial \mathbb{E}_{t-1}\pi_t} = -\frac{\beta}{\kappa}$$

from which the results follow using Propositions 5 and 7.

A.11 Proof of Proposition 11

The proof follows the same steps as in Proposition 7. The envelope condition is the same, given that the additional term $-c(\pi_t - \mathbb{E}_{t-1}[\pi_t|\tilde{\theta}_t])$ in V_t depends on reported types and not true types. From here, the value function at date t under our proposed mechanism given by

$$\mathcal{W}_t(\theta^t) = -c(\pi_t - \mathbb{E}_{t-1}\pi_t) + V_t + \beta\mathbb{E}_t\left[\mathcal{W}_t(\theta^{t+1})|\theta_t\right]$$
$$= -(c + b_{t-1})(\pi_t - \mathbb{E}_{t-1}\pi_t) + U_t + \beta\mathbb{E}_t\left[\mathcal{W}_t(\theta^{t+1})|\theta_t\right]$$
$$= -\nu_{t-1}(\pi_t - \mathbb{E}_{t-1}\pi_t) + U_t + \beta\mathbb{E}_t\left[\mathcal{W}_t(\theta^{t+1})|\theta_t\right]$$

which is the same value function as in the proof of Proposition 7 when evaluated at the constrained efficient allocation. Thus the result follows using the same proof as for Proposition 7.

A.12 Proof of Proposition 12

See Appendix B for the detailed setup and proof.

A.13 Proof of Corollary 13

The proof follows immediately from the definition of Γ_t, which is equal to zero if $\theta_t \in \{\underline{\theta}, \overline{\theta}\}$. When $\Gamma_t = 0$, the allocation rule is constrained efficient for all Γ_{t+k}, $k \geq 1$, so the optimal mechanism reverts to constrained efficiency, which is implemented by the dynamic inflation target.

B Second-Best Mechanisms with Costly Transfers

In this extension, we provide the details behind the derivations of the second-best allocation with costly transfers, including proofs.

We begin by characterizing local incentive compatibility from the Envelope Theorem, in order to solve the problem using a first-order approach that relaxes the required monotonicity constraint.[35] The Bellman representation of the central bank's problem is

$$\mathcal{W}_t\left(\theta^t\right) = \sup_{\tilde{\theta}} \left\{U_t + T_t + \beta\mathbb{E}_t\left[\mathcal{W}_{t+1}|\theta_t\right]\right\} \tag{23}$$

As before, π_t, T_t are functions of the history of reported types, not true types, while $\mathcal{W}_t$ is a function of both past *reported* types and the current *true* type.

[35] See Pavan et al. (2014).

We adopt the notational convention that $\frac{\partial U_t}{\partial \theta_t}$ is the derivative of U_t in the true type θ_t, while $\frac{\partial U_t}{\partial \tilde{\theta}_t}$ is the derivative in the reported type. In equilibrium, the reported type and the true type coincide, but the derivatives are distinct. Using the envelope condition derived in the first part of the proof of Proposition 7, we obtain an integral incentive compatibility condition

$$\mathcal{W}_t(\theta^t) = \int_{\underline{\theta}}^{\theta_t} \frac{\partial U_t(\theta^{t-1}, s_t)}{\partial s_t} ds_t + \beta \int_{\underline{\theta}}^{\theta_t} \mathbb{E}_t \left[\mathcal{W}_{t+1} \frac{\partial f_t(\theta_{t+1}|s_t)/\partial s_t}{f_t(\theta_{t+1}|s_t)} |s_t \right] ds_t \tag{24}$$

Integral incentive compatibility relates the total date-t utility to the central bank to two information rents. Note that due to shock persistence, the central bank earns information rents not only due to the effect on current flow utility, but also on the conditional probability distribution.[36]

Integral incentive compatibility (24) gives a Bellman representation to the value function, in terms of only the allocation rule. We can re-express this Bellman equation in sequence form by iterating the Bellman equation forward. Doing so, we obtain the following result characterizing this sequence representation.

Lemma 14. *The value function $\mathcal{W}_t$ can be represented as*

$$\mathcal{W}_t(\theta^t) = \mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s B_t^s(\theta^{t+s}) \middle| \theta_t \right] \quad \forall t,$$

where B_t^s is given by

$$B_t^s(\theta^{t+s}) = \prod_{k=0}^{s-1} \frac{1}{f_{t+k}(\theta_{t+k+1}|\theta_{t+k})} \times \int_{s_t \le \theta_t, \dots, s_{t+s} \le \theta_{t+s}} \frac{\partial U_{t+s}(\theta^{t-1}, s_t, \dots, s_{t+s})}{\partial s_{t+s}} \prod_{k=0}^{s-1} \frac{\partial f_{t+k}(\theta_{t+k+1}|s_{t+k})}{\partial s_{t+k}} ds_{t+s} \dots ds_t.$$

Lemma 14 allows us to represent the principal's optimization problem in a tractable way. Given an allocation rule for inflation, we use the characterization of the value function in Lemma 14 as well as the Bellman equation to characterize the transfer rule

[36] Without loss of generality, we have set the constant of integration $\mathcal{W}_t(\theta^{t-1}, \underline{\theta}) = 0$ and have used the outside option normalization $\underline{W}_0(\underline{\theta}) = 0$. It is without loss to set the integration constant to zero, since it can always be incorporated into the prior period's value function.

which implements the allocation,

$$T_t = \mathcal{W}_t - U_t - \beta \mathbb{E}_t[\mathcal{W}_{t+1}|\theta_t].$$

We can then substitute the implementing taxes into the government's utility function, and obtain the following result characterizing the relaxed social planning problem.

Lemma 15. *The relaxed social planning problem can be written as*

$$\max_{\{\pi_t\}} \mathbb{E}_{-1}\left[\sum_{t=0}^{\infty} \beta^t \left[-\frac{\kappa}{1+\kappa} B_0^t + U_t\right]\right],$$

where B_0^t is given as in Lemma 14. The implementing transfer rule is given by

$$T_t = \mathcal{W}_t - U_t - \beta \mathbb{E}_t[\mathcal{W}_{t+1}|\theta_t],$$

where $\mathcal{W}_t$ is given as a function of the allocation rule as in Lemma 14.

Lemma 15 provides a characterization of the relaxed social planning problem, subject to integral incentive compatibility. From this, we can characterize the optimal allocation in Proposition 12.[37]

B.1 Second best with Average Transfers

In the baseline model, we impose the assumption that the outside option takes the form $\mathcal{W}_0(\theta^0) \geq 0$. We might alternatively have expressed this in the form

$$\int_{\theta_0} \mathcal{W}_0(\theta^0) f(\theta_0|\theta_{-1}) d\theta_0 \geq 0$$

The core difference between these two assumptions from a modeling perspective is on the timing of information arrival versus the participation decision. Under the baseline assumption, either θ_0 is already known to the central bank, or the central bank has the opportunity to revert to the outside option after learning θ_0. Under the second assumption, θ_0 is not known to the central bank, and the central bank does not have the option to revert to the outside option after learning it.

[37] We characterize the optimal allocation assuming that π_t is interior.

Under this alternative structure, the optimality of the dynamic inflation target returns. In particular, implementable allocations are still defined as in Lemma 14, while the transfer rule is $T_t(\theta^t) = \mathcal{W}_t - U_t - \beta\mathbb{E}_t\left[\mathcal{W}_{t+1}|\theta_t\right]$. The average participation constraint implies that we have

$$0 = E_{-1}\mathcal{W}_0 = \mathbb{E}_{-1}\sum_{t=0}^{\infty}\beta^t(U_t + T_t),$$

which is markedly different from the baseline model. In particular, substituting this expression into social welfare, we obtain the social optimization problem

$$\max_{\{\pi_t\}} \mathbb{E}_{-1}\sum_{t=0}^{\infty}\beta^t(1+\kappa)U_t$$

implying that the optimal allocation rule is constrained efficient. From here, we obtain the optimality of the dynamic inflation target.

Proposition 16. *Suppose that the participation constraint takes the form*

$$\int_{\theta_0}\mathcal{W}_0(\theta^0)f(\theta_0|\theta_{-1})d\theta_0 \geq 0$$

Then, the optimal mechanism is a dynamic inflation target, and yields the constrained efficient allocation.

The intuition behind Proposition 12 is straight-forward: under the average constraint, the government can capture the full social surplus and simply reduce the average transfer to the central bank at date 0 to satisfy the participation constraint. This implies that the government chooses the mechanism and allocation that maximize social surplus, which is the dynamic inflation target.

B.2 Proofs

B.2.1 Proof of Lemma 14

Suppose that we take the Bellman equation:

$$\mathcal{W}_t(\theta^t) = \int_{\underline{\theta}}^{\theta_t}\frac{\partial U_t(\theta^{t-1}, s_t)}{\partial s_t}ds_t + \beta\int_{\underline{\theta}}^{\theta_t}E_t\left[\mathcal{W}_{t+1}\frac{\partial f_t(\theta_{t+1}|s_t)/\partial s_t}{f_t(\theta_{t+1}|s_t)}|s_t\right]$$

And iterate it forward once. Iterating forward once, we obtain:

$$\mathcal{W}_t(\theta^t) = \int_{\underline{\theta}}^{\theta_t} E_t\left[\frac{\partial U_t(\theta^{t-1}, s_t)}{\partial s_t} ds_t + \frac{\partial f_t\left(\theta_{t+1}|s_t\right)/\partial s_t}{f_t\left(\theta_{t+1}|s_t\right)}\beta\left[\int_{\underline{\theta}}^{\theta_{t+1}} \frac{\partial U_t(\theta^{t-1}, s_t, s_{t+1})}{\partial s_{t+1}} + E_{t+1}\mathcal{W}_{t+2}\frac{f_{t+1}(\theta_{t+2}|s}{f_{t+1}(\theta_{t+}}\right.\right.$$

Iterating forward, suppose that we define the following recursive operator. In particular, we define:

$$\mathcal{B}_t^0(g,\theta) = \int_{\underline{\theta}}^{\theta} g ds_t$$

Note that for the function $g_t^0 = \frac{\partial U_t(\theta^{t-1}}{\partial s_t}$, we have that $\mathcal{B}_t^0$ is the first term in the infinite series defining $\mathcal{W}_t$.

And suppose we define next:

$$\mathcal{B}_t^1\left(g,\theta\right) = \int_{\underline{\theta}}^{\theta} E_t\left[\frac{\partial f_t(\theta_{t+1}|s_t)/\partial s_t}{f_t(\theta_{t+1}|s_t)} g\middle| s_t\right] ds_t$$

Consider the function $g_t^1 = \int_{\underline{\theta}}^{\theta_{t+1}} \frac{\partial U_{t+1}(\theta^{t-1}, s_t, s_{t+1})}{\partial s_{t+1}} ds_{t+1}$. Taking the function $\mathcal{B}_t^1\left(g_t^1, \theta_t\right)$ and multiplying by β, we obtain the second term in the infinite series for $\mathcal{W}_t$.

From here, we define a recursive operator. Consider a function g_t^s that is a date $t+s$ adapted function. We define the operator:

$$\mathcal{B}_t^2\left(g_t^2, \theta_t\right) = \mathcal{B}_t^1\left(\mathcal{B}_{t+1}^1\left(g_t^2, \theta_{t+1}\right), \theta_t\right)$$

So that we have:

$$\mathcal{B}_t^2\left(g_t^2, \theta_t\right) = \int_{\underline{\theta}}^{\theta_t} E_t\left[\frac{\partial f_t(\theta_{t+1}|s_t)/\partial s_t}{f_t(\theta_{t+1}|s_t)} \int_{\underline{\theta}}^{\theta_{t+1}} E_{t+1}\left[\frac{\partial f_{t+1}(\theta_{t+2}|s_{t+1})/\partial s_{t+1}}{f_{t+1}(\theta_{t+2}|s_{t+1})} g_t^2\left(s_{t+1}, \theta_{t+2}\right)\middle| s_{t+1}\right] ds_{t+1}\middle| s_t\right]$$

Which, when $g_t^2(s_t, s_{t+1}, \theta_{t+2}) = \int_{\underline{\theta}}^{\theta_{t+2}} \frac{\partial U_{t+2}\left(\theta^{t-1}, s_t, s_{t+1}, s_{t+2}\right)}{\partial s_{t+2}} ds_{t+2}$ and multiplied by β^2, gives us the next term in the infinite series defining $\mathcal{W}_t$.

Continuosly defining these recursive operators as such, and defining functions $g_t^s\left(s_t, ..., s_{t+s-1}, \theta_{t+s}\right) = \int_{\underline{\theta}}^{\theta_{t+s}} \frac{\partial U_{t+s}\left(\theta^{t-1}, s_t, ..., s_{t+s}\right)}{\partial s_{t+s}}$, we obtain the infinite series that characterizes $\mathcal{W}_t$.

In other words, we can construct such recursive operators. From here, we look to simplify these operators. Let us start from the operator $\mathcal{B}_t^1\left(g, \theta_t\right)$. In particular, we have:

$$
\begin{aligned}
\mathcal{B}_t^1(g,\theta_t) &= \int_{\underline{\theta}}^{\theta_t} E_t\left[\frac{\partial f_t(\theta_{t+1}|s_t)/\partial s_t}{f_t(\theta_{t+1}|s_t)} g(s_t,\theta_{t+1})\right| s_t\bigg] ds_t \\
&= \int_{\underline{\theta}}^{\theta_t}\int_{\theta_{t+1}} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} g(s_t,\theta_{t+1}) d\theta_{t+1} ds_t \\
&= \int_{\theta_{t+1}}\left[\int_{\underline{\theta}}^{\theta_t} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} g(s_t,\theta_{t+1}) ds_t\right] d\theta_{t+1} \\
&= \int_{\theta_{t+1}} \frac{\left[\int_{\underline{\theta}}^{\theta_t} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} g(s_t,\theta_{t+1}) ds_t\right]}{f_t(\theta_{t+1}|\theta_t)} f_t(\theta_{t+1}|\theta_t) d\theta_{t+1} \\
&= E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} g(s_t,\theta_{t+1}) ds_t\right]\right| \theta_t\bigg]
\end{aligned}
$$

In particular, as applied to the function $g_t^1 = \int_{\underline{\theta}}^{\theta_{t+1}} \frac{\partial U_{t+1}(\theta^{t-1},s_t,s_{t+1})}{\partial s_{t+1}} ds_{t+1}$, we obtain:

$$
\mathcal{B}_t^1(g,\theta_t) = E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t}\int_{\underline{\theta}}^{\theta_{t+1}} \frac{\partial U_{t+1}(\theta^{t-1},s_t,s_{t+1})}{\partial s_{t+1}} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} ds_{t+1} ds_t\right]\right| \theta_t\bigg]
$$

Which is of the form in the Lemma.

Now, let us consider the second operator. We have:

$$
\mathcal{B}_t^2(g,\theta_t) = \mathcal{B}_t^1\left(\mathcal{B}_{t+1}^1(g,\theta_{t+1}),\theta_t\right)
$$

Recall that the simplified operator above expresses:

$$
\mathcal{B}_t^1(g,\theta_t) = E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t} \frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t} g(s_t,\theta_{t+1}) ds_t\right]\right| \theta_t\bigg]
$$

In other words, we have along history (θ^{t-1}, s_t):

$$
\mathcal{B}_{t+1}^1(g,\theta_{t+1}) = E_{t+1}\left[\frac{1}{f_{t+1}(\theta_{t+2}|\theta_{t+1})}\left[\int_{\underline{\theta}}^{\theta_{t+1}} \frac{\partial f_{t+1}(\theta_{t+2}|s_{t+1})}{\partial s_{t+1}} g(s_t,s_{t+1},\theta_{t+2}) ds_{t+1}\right]\right| \theta_{t+1}\bigg]
$$

And applying this into the operator defining $\mathcal{B}_t^2$, we obtain:

$$
\begin{aligned}
\mathcal{B}_t^2\left(g,\theta_t\right) &= E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t}\frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t}\mathcal{B}_{t+1}^1\left(g,\theta_{t+1}\right)ds_t\right]\middle|\,\theta_t\right] \\
&= E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t}\frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t}E_{t+1}\left[\frac{1}{f_{t+1}(\theta_{t+2}|\theta_{t+1})}\left[\int_{\underline{\theta}}^{\theta_{t+1}}\frac{\partial f_{t+1}(\theta_{t+2}|s_{t+1})}{\partial s_{t+1}}g(s_t,s_{t+1},\theta_{t+2}\right.\right.\right.\right. \\
&= E_tE_{t+1}\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\left[\int_{\underline{\theta}}^{\theta_t}\frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t}\left[\frac{1}{f_{t+1}(\theta_{t+2}|\theta_{t+1})}\left[\int_{\underline{\theta}}^{\theta_{t+1}}\frac{\partial f_{t+1}(\theta_{t+2}|s_{t+1})}{\partial s_{t+1}}g(s_t,s_{t+1},\theta_{t+2}\right.\right.\right.\right. \\
&\overbrace{=}^{\text{LIE}} E_t\left[\frac{1}{f_t(\theta_{t+1}|\theta_t)}\frac{1}{f_{t+1}(\theta_{t+2}|\theta_{t+1})}\left[\int_{\underline{\theta}}^{\theta_t}\int_{\underline{\theta}}^{\theta_{t+1}}\frac{\partial f_t(\theta_{t+1}|s_t)}{\partial s_t}\frac{\partial f_{t+1}(\theta_{t+2}|s_{t+1})}{\partial s_{t+1}}g(s_t,s_{t+1},\theta_{t+2})ds_{t+1}\right.\right.
\end{aligned}
$$

And substituting in $g_t^2 = \int_{\underline{\theta}}^{\theta_{t+2}}\frac{\partial U_{t+2}\left(\theta^{t-1},s_t,s_{t+1},s_{t+2}\right)}{\partial s_{t+2}}ds_{t+2}$, we get the next expression from the Lemma. From here, the result follows from repeated iteration.

B.2.2 Proof of Lemma 15

For any allocation rule, T_t provides the implementation. Recall that the government's welfare is given by:

$$\max E_{-1}\left[\sum_{t=0}^{\infty}\beta^t U_t - \kappa T_t\right],$$

Recall that bank welfare is given by:

$$\mathcal{W}_0 = E_0\sum_{t=0}^{\infty}\left[\beta^t U_t + T_t\right]$$

In other words, we always have:

$$-E_0\sum_{t=0}^{\infty}T_t = E_0\sum_{t=0}^{\infty}\beta^t U_t - \mathcal{W}_0$$

Substituting in above, by Law of Iterated Expectations we obtain the planning problem:

$$\max E_{-1}\left[-\kappa\mathcal{W}_0 + \sum_{t=0}^{\infty}\beta^t(1+\kappa)U_t\right],$$

and where lastly, we use Lemma 4 substitute in for $\mathcal{W}_0$ to obtain the result.

B.3 Proof of Proposition 12

Recall that our objective function for the second-best optimization problem was given by:

$$\max \int_{\theta_0} \left[\sum_{t=0}^{\infty} \beta^t \left[-\frac{\kappa}{1+\kappa} \mathcal{B}_0^s \left(g_0^t, \theta_0 \right) + U_t \left(\pi_t, \pi_{t+1}, \theta_t, \theta_t \right) \right] \right] dF_0(\theta_0)$$

Note that given the optimal mechanism implements truthful reporting, we may substitute in $\tilde{\theta}_t = \theta_t$.

Recall further the simplified form of the operators:

$$\mathcal{B}_t^s = E_t \left[\prod_{k=0}^{s-1} \frac{1}{f_{t+k}(\theta_{t+k+1}|\theta_{t+k})} \int_{s_t \le \theta_t, \ldots, s_{t+s} \le \theta_{t+s}} \frac{\partial U_{t+s}(\theta^{t-1}, s_t, \ldots, s_{t+s})}{\partial s_{t+s}} \prod_{k=0}^{s-1} \frac{\partial f_{t+k}(\theta_{t+k+1}|s_{t+k})}{\partial s_{t+k}} ds_{t+s} \ldots ds_t \right| \theta$$

Now, denote the *realized value* of the operator $\mathcal{B}_0^t$ by:

$$B_0^t(\theta^t) = \prod_{k=0}^{t-1} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \int_{s_0 \le \theta_0, \ldots, s_t \le \theta_t} \frac{\partial U_t(s_0, \ldots, s_t)}{\partial s_t} \prod_{k=0}^{t-1} \frac{\partial f_k(\theta_{k+1}|s_k)}{\partial s_k} ds_t \ldots ds_0$$

So that $B_0^t(\theta^t)$ is a random variable derived from the history θ^t of shocks. Given the definition of this random variable, denote E_{-1} to be the beginning-of-period-0 expectation, not conditional on the information θ_0. From here, we can rewrite the objective function of the government as:

$$\max E_{-1} \left[\sum_{t=0}^{\infty} \beta^t \left[-\frac{\kappa}{1+\kappa} B_0^t(\pi_t, \pi_{t+1}, \theta_t|\theta^{t-1}) + (1+\kappa) U_t(\pi_t, \pi_{t+1}, \theta_t|) \right] \right]$$

From here, consider the optimal choice of inflation $\pi_t(z^t)$, for a realized history $\theta^t = z^t$ of shocks. Note that the solution can be written in the form (for $t \ge 1$):

$$\frac{\partial U_{t-1}}{\partial \pi_t(z^t)} f(z^{t-1}) + \beta \frac{\partial U_t}{\partial \pi_t(z^t)} f(z^t) = \frac{\kappa}{1+\kappa} E_{-1} \sum_{s=t-1}^{t} \beta^{s-(t-1)} \frac{d}{d\pi_t(z^t)} B_0^s(\pi_s, \pi_{s+1}, \theta_s|\theta^s)$$

So that all that remains is to characterize the derivatives of B_0^s with respect to $\pi_t(z^t)$. When $s = t$, we have:

$$\frac{d}{dz^t} B_0^t(\theta^t) = \frac{d}{\pi_t(z^t)} \left[\prod_{k=0}^{t-1} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \int_{s_0 \le \theta_0, \ldots, s_t \le \theta_t} \frac{\partial U_t(s_0, \ldots, s_t)}{\partial s_t} \prod_{k=0}^{t-1} \frac{\partial f_k(\theta_{k+1}|s_k)}{\partial s_k} ds_t \ldots ds_0 \right]$$

Note that $\pi_t(z^t)$ appears in $\frac{\partial U_t(s_0,\ldots,s_t)}{\partial s_t}$ only along the path given by $s_0 = z_0$, $s_1 = z_1$, ... , $s_t = z_t$. Essentially then, this derivative at a single point $\pi_t(z^t)$ comes down to extracting the derivative along that path under the integral. The derivative along that path is then given by:

$$\frac{d}{dz^t} B_0^t(\theta^t) = \mathbf{1}_{z_0 \leq \theta_0, \ldots, z_t \leq \theta_t} \prod_{k=0}^{t-1} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \frac{\partial^2 U_t}{\partial z_t \partial \pi_t(z^t)} \prod_{k=0}^{t-1} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k}$$

Note the subtlety that the θ's are preserved, as the realization of the random history, whereas the s's are replaced by z's, as the path under the integrals that leads to the history z^t under the integrals. It is worth remembering then, when we substitute into the expectation, that θ_t is a random variable, and z^t is (fixed) the history being differentiated along, and so is not a random variable.

Note that by exactly the same logic, we obtain $\forall t \geq 2$

$$\frac{d}{dz^t} B_0^{t-1}(\theta^{t-1}) = \mathbf{1}_{z_0 \leq \theta_0, \ldots, z_{t-1} \leq \theta_{t-1}} \prod_{k=0}^{t-2} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \frac{\partial^2 U_{t-1}}{\partial z_{t-1} \partial \pi_t(z^t)} \prod_{k=0}^{t-2} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k}$$

As a result, the right-hand side of the first-order condition becomes $\forall t \geq 2$

$$\begin{aligned}
\frac{1+\kappa}{\kappa}\text{RHS} &= E_{-1} \sum_{s=t-1}^{t} \frac{d}{d\pi_t(z^t)} B_0^s(\pi_s, \pi_{s+1}, \theta_s|\theta^s) \\
&= E_{-1} \left[\mathbf{1}_{z_0 \leq \theta_0, \ldots, z_{t-1} \leq \theta_{t-1}} \prod_{k=0}^{t-2} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \frac{\partial^2 U_{t-1}}{\partial z_{t-1} \partial \pi_t(z^t)} \prod_{k=0}^{t-2} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k} \right] \\
&\quad + \beta E_{-1} \left[\mathbf{1}_{z_0 \leq \theta_0, \ldots, z_t \leq \theta_t} \prod_{k=0}^{t-1} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \frac{\partial^2 U_t}{\partial z_t \partial \pi_t(z^t)} \prod_{k=0}^{t-1} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k} \right] \\
&= \frac{\partial^2 U_{t-1}}{\partial z_{t-1} \partial \pi_t(z^t)} E_{-1} \left[\mathbf{1}_{z_0 \leq \theta_0, \ldots, z_{t-1} \leq \theta_{t-1}} \prod_{k=0}^{t-2} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k} \right] \\
&\quad + \frac{\partial^2 U_t}{\partial z_t \partial \pi_t(z^t)} \beta E_{-1} \left[\mathbf{1}_{z_0 \leq \theta_0, \ldots, z_t \leq \theta_t} \prod_{k=0}^{t-1} \frac{1}{f_k(\theta_{k+1}|\theta_k)} \prod_{k=0}^{t-1} \frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k} \right]
\end{aligned}$$

Where here, we applied the fact that we have chosen a specific history z^t, so that the cross-partials above are *not* random variables, but rather are specific realizations of those random variables. By contrast, the part inside the expectation corresponds to histories which contain these specific histories, and so are random variables.

Now, consider these two expectations. Now, we define $\Omega_t(z^t)$ by:

$$
\begin{aligned}
\Omega_t(z^t) &\equiv E_{-1}\left[\mathbf{1}_{z_0\le\theta_0,\dots,z_t\le\theta_t}\prod_{k=0}^{t-1}\frac{1}{f_k(\theta_{k+1}|\theta_k)}\prod_{k=0}^{t-1}\frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k}\right] \\
&= \int_{z_t}^{\overline{\theta}}\int_{z_{t-1}}^{\overline{\theta}}\dots\int_{z_0}^{\overline{\theta}}\prod_{k=0}^{t-1}\frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k}f(\theta_0)d\theta_t\dots d\theta_0 \\
&= \int_{z_t}^{\overline{\theta}}\frac{\partial f_k(\theta_t|z_{t-1})}{\partial z_k}\left[\int_{z_{t-1}}^{\overline{\theta}}\dots\int_{z_0}^{\overline{\theta}}\prod_{k=0}^{t-1}\frac{\partial f_k(\theta_{k+1}|z_k)}{\partial z_k}f(\theta_0)d\theta_{t-1}\dots d\theta_0\right]d\theta_t \\
&= \int_{z_t}^{\overline{\theta}}\frac{\partial f_k(\theta_t|z_{t-1})}{\partial z_{t-1}}\Omega_{t-1}(z^{t-1})d\theta_t \\
&= \Omega_{t-1}\left(z^{t-1}\right)\int_{z_t}^{\overline{\theta}}\frac{\partial f_k(\theta_t|z_{t-1})}{\partial z_{t-1}}d\theta_t
\end{aligned}
$$

Which is well-defined for all $t \ge 1$. However, it requires an initial condition $\Omega_0(z^0)$. It is helpful to define this initial condition in the date 1 FOC. Note that at date 1, we have:

$$
\mathcal{B}_0^{t-1}(\theta^{t-1}) = \mathcal{B}_0^0(\theta^0) = \int_{\underline{\theta}}^{\theta_0}\frac{\partial U_0}{\partial s_0}ds_0
$$

So that we have $\frac{d}{d\pi_t(z^t)}\mathcal{B}_0^{t-1}(\theta^{t-1}) = \mathbf{1}_{z_0\le\theta_0}\frac{\partial U_0}{\partial\pi_1(z^1)}$. In particular then, the expectation is simply:

$$
E_{-1}\left[\mathbf{1}_{z_0\le\theta_0}\right] = \int_{z_0}^{\overline{\theta}}f(\theta_0)d\theta_0 = 1 - F(z_0)
$$

So that we have initial condition $\Omega_0(z^0) = 1 - F(z_0)$.

This gives us a state space reduction property, where we can fully determine Ω_t from Ω_{t-1} and z_{t-1} by a recursive sequence, where the initial value is $\Omega_0(z^0) = 1 - F(z_0)$.

From here, we can substitute back into the FOCs:

$$
(1+\kappa)\left[\frac{\partial U_{t-1}}{\partial\pi_t(z^t)}f(z^{t-1}) + \beta\frac{\partial U_t}{\partial\pi_t(z^t)}f(z^t)\right] = \kappa\left[\Omega_{t-1}(z^{t-1})\frac{\partial^2 U_{t-1}}{\partial z_{t-1}\partial\pi_t(z^t)} + \beta\Omega_t(z^t)\frac{\partial^2 U_t}{\partial z_t\partial\pi_t(z^t)}\right]
$$

From here, it is helpful to divide through by $f(z^{t-1})$:

$$
(1+\kappa)\left[\frac{\partial U_{t-1}}{\partial\pi_t(z^t)} + \beta\frac{\partial U_t}{\partial\pi_t(z^t)}f(z_t|z_{t-1})\right] = \kappa\left[\frac{\Omega_{t-1}(z^{t-1})}{f(z^{t-1})}\frac{\partial^2 U_{t-1}}{\partial z_{t-1}\partial\pi_t(z^t)} + \beta\frac{\Omega_t(z^t)}{f(z^t)}\frac{\partial^2 U_t}{\partial z_t\partial\pi_t(z^t)}f(z_t|z_{t-1})\right]
$$

And from here, we define $\Gamma_t(z^t) = \frac{\Omega_t(z^t)}{f(z^t)}$. Note that we have:

$$\Gamma_t(z^t) = \frac{\Omega_t(z^t)}{f(z^t)} = \frac{\Omega_{t-1}(z^{t-1})}{f(z^t)} \frac{\int_{z_t}^{\overline{\theta}} \frac{\partial f_k(\theta_t|z_{t-1})}{\partial z_k} d\theta_t}{f(z_t|z_{t-1})} = \Gamma_{t-1}(z^{t-1}) \frac{\int_{z_t}^{\overline{\theta}} \frac{\partial f_k(\theta_t|z_{t-1})}{\partial z_k} d\theta_t}{f(z_t|z_{t-1})}$$

Giving us our key result for $t \geq 1$.

Note that the relevant initial condition is $\Gamma_0 = \frac{1-F(z_0)}{f(z_0)}$. This is the standard term in evaluating the virtual value in static mechanism design problems, and it is not surprising that it appears here. What it notable is that this term appears in the *date 1* optimality condition, in addition (as we will see) to the date-0 one. This is because of the time consistency problem.

Lastly, we can evaluate the FOC for π_0. In π_0, there is no time consistency element, and we are left with the simple tradeoff between current π and transfers. Repeating the steps from above, we obtain the simple condition

$$\frac{\partial U_0}{\partial \pi_0} = \frac{\kappa}{1+\kappa} \Gamma_0(z^0) \frac{\partial^2 U_0}{\partial z_0 \partial \pi_0}$$

which is a standard virtual value condition. This gives the full result.

B.3.1 Proof of Proposition 16

The optimization problem

$$\max_{\{\pi_t\}} E_{-1} \sum_{t=0}^{\infty} \beta^t (1+\kappa) U_t$$

yields the constrained efficient allocation rule. The result of Proposition 7 gives the mechanism that implements this allocation rule, giving the result.