NBER WORKING PAPER SERIES

OPTIMAL MONETARY POLICY WITH HETEROGENEOUS AGENTS: DISCRETION, COMMITMENT, AND TIMELESS POLICY

Eduardo Dávila
Andreas Schaab

Working Paper 30961
http://www.nber.org/papers/w30961

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2023

Nothing to disclose. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w30961

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Eduardo Dávila and Andreas Schaab. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Optimal Monetary Policy with Heterogeneous Agents: Discretion, Commitment, and Timeless Policy
Eduardo Dávila and Andreas Schaab
NBER Working Paper No. 30961
February 2023
JEL No. E52,E61

ABSTRACT

This paper characterizes optimal monetary policy in a canonical heterogeneous-agent New Keynesian (HANK) model with wage rigidity. Under discretion, a utilitarian planner faces the incentive to redistribute towards indebted, high marginal utility households, which is a new source of inflationary bias. With commitment, i) zero inflation is the optimal long-run policy, ii) time-consistent policy requires both inflation and distributional penalties, and iii) the planner trades off aggregate stabilization against distributional considerations, so Divine Coincidence fails. We compute optimal stabilization policy in response to productivity, demand, and cost-push shocks using sequence-space methods, which we extend to Ramsey problems and welfare analysis.

Eduardo Dávila
Yale University
Department of Economics
28 Hillhouse Avenue, Room 304
New Haven, CT 06520-8268
and NBER
eduardo.davila@yale.edu

Andreas Schaab
Toulouse School of Economics
andreas.schaab@tse-fr.eu

An online appendix is available at http://www.nber.org/data-appendix/w30961

1 Introduction

There is large heterogeneity in households' exposure to business cycle fluctuations. At the same time, there is now a growing consensus that monetary policy has distributional consequences—a view supported by mounting empirical evidence (Doepke and Schneider, 2006; Coibion et al., 2017; Ampudia et al., 2018) and the burgeoning heterogeneous-agent New Keynesian (HANK) literature (McKay et al., 2016; Kaplan et al., 2018; Auclert, 2019; Auclert et al., 2018). Household heterogeneity may therefore be an important determinant of the welfare impact of monetary policy and should inform the study of optimal policy design. However, accounting for rich heterogeneity and incomplete markets in dynamic optimal policy problems remains challenging.

In this paper, we characterize optimal monetary policy in a canonical one-asset HANK economy with wage rigidity, which represents a minimal departure from the representative-agent New Keynesian (RANK) model. Our goal is to systematically revisit the canonical New Keynesian consensus on optimal monetary policy (Clarida et al., 1999; Woodford, 2003; Galí, 2015). To do so, we structure our analysis of optimal monetary policy to parallel that of Clarida et al. (1999), starting with policy under discretion in Section 3 and studying optimal policy under commitment in Section 4. Concluding with a quantitative analysis in Section 5, we compute optimal monetary policy both non-linearly and using sequence-space perturbation methods (Boppart et al., 2018; Auclert et al., 2021), which we extend to Ramsey problems and welfare analysis.

Optimal monetary policy under discretion. Under discretion, a utilitarian planner in a HANK economy has an incentive to raise output above natural output and overheat the economy, even in the absence of markup distortions. This occurs because the planner values redistribution toward indebted, high marginal utility households via lower interest rates. At the optimum, the planner trades off this novel redistribution motive against aggregate stabilization. However, when agents anticipate the planner's incentives to lower interest rates, inflationary bias in the sense of Barro and Gordon (1983) emerges in equilibrium. Quantitatively, the redistribution motive dominates the standard markup distortion as a source of inflationary bias. The gains from commitment are consequently larger in heterogeneous-agent economies.

Optimal monetary policy with commitment. Motivated by the results under discretion, we study optimal policy under commitment in three steps. Each step isolates an important dimension of optimal monetary policy design: long-run policy, time consistency, and stabilization policy. We study the implications of household heterogeneity for optimal monetary policy along each of these dimensions.

In the first step, after introducing the standard Ramsey problem and characterizing the associated Ramsey plan, we study optimal long-run policy. We show that the optimal stationary equilibrium under commitment features zero inflation, eliminating the inflationary bias of policy under discretion in the long run. This result is due to the fact that inflation and the nominal interest rate affect households' financial income symmetrically, which can be seen as a relevant benchmark. Therefore, since the long-run real interest rate is invariant to policy, but inflation is costly while adjusting the nominal rate is not, the planner finds it optimal to keep inflation at zero in the stationary Ramsey plan.

In the second step, we show that while the standard Ramsey problem eliminates inflationary bias in the long run, it still suffers from inflationary bias in the short run. This is a due to a "time-0 problem" (Kydland and Prescott, 1980) associated with two dimensions of time inconsistency. The first source of time inconsistency is the forward-looking Phillips curve, through which inflation expectations enter the Ramsey problem. This time consistency problem has been widely studied in RANK economies by the literature following Barro and Gordon (1983). In the presence of household heterogeneity, a new second time consistency problem emerges because forward-looking individual Bellman equations appear as constraints in the Ramsey problem.

While the standard Ramsey planner chooses policy with commitment from time 0 onwards, time inconsistency still manifests at time 0. In order to find a "timeless" planning solution, we extend the approach of Marcet and Marimon (2019) to our setting (i.e., continuous-time heterogeneous-agent economies) by introducing timeless penalties for each forward-looking implementability condition. We then define a timeless Ramsey problem, which augments the standard Ramsey problem with such timeless penalties, and prove that it no longer suffers from a time-0 problem: the planner has no incentive to deviate from the stationary Ramsey plan in the absence of shocks. Hence, the timeless Ramsey problem eliminates inflationary bias in both the short and the long run.

We analytically characterize the two timeless penalties required by the timeless Ramsey problem: an inflation penalty and a distributional penalty. We first show that the inflation penalty, which is already present in RANK economies, depends on novel distributional considerations in HANK. When households are heterogeneous, changes in aggregate economic activity have distributional consequences. The standard Ramsey planner's incentive to generate inflation out of steady state at time 0, which the inflation penalty is designed to counteract, is consequently also governed by distributional considerations. Second, we show that the new distributional penalty penalizes the welfare gains of indebted, high marginal utility households. While it may seem counterintuitive that a utilitarian planner penalizes high marginal utility households, this is precisely to counteract the planner's time inconsistent incentive to redistribute towards such households, which becomes a source of inflationary bias under discretion. Finally, we show that the distributional penalty solves a novel promise-keeping Kolmogorov forward equation.

Concluding our discussion of time consistency, we explore whether a central bank that sets policy under discretion can still implement the optimal commitment solution under an appropriate institutional arrangement or with the appropriate penalties or targets. We first show that the timeless Ramsey plan can be implemented under discretion if the planner's objective is augmented to incorporate the appropriate time-varying inflation and distributional penalties. Moreover, a strict zero-inflation target implements the timeless Ramsey plan in the absence of shocks, while a modified flexible inflation target around zero inflation can implement optimal stabilization policy under commitment.

In the third and final step, we study optimal stabilization policy under the timeless Ramsey problem, which allows us to separate the pure stabilization motive from the time-0 problem. We characterize an analytical targeting rule for optimal stabilization policy in response to demand, supply, and cost-push shocks, and use it to illustrate the departures from optimal policy in RANK, which it nests. In a RANK economy, no tradeoff emerges between inflation and output in the absence of cost-push shocks; the planner finds it optimal to simultaneously close both the inflation and output gaps. In HANK economies, on the other hand, this Divine Coincidence result generically fails even in the absence of cost-push shocks. The planner now accounts for the distributional impact of policies and perceives a tradeoff between aggregate stabilization and distributional considerations. That is, even in the absence of cost-push shocks and with the appropriate employment subsidy, the planner finds it optimal not to simultaneously close the inflation and output gaps in response to shocks.

Quantitative analysis in sequence space. This paper extends the sequence-space approach (Boppart et al., 2018; Auclert et al., 2021) to Ramsey problems and welfare analysis. We develop a general sequence-space representation of (timeless) Ramsey plans that builds on Auclert et al. (2021)'s sequence-space representation of competitive equilibrium in heterogeneous-agent economies. Under this representation, a Ramsey plan is a system of equations that take as inputs the time paths of aggregate allocations and prices, aggregate multipliers, policies, and shocks.

While our approach allows us to characterize and compute timeless Ramsey plans non-linearly, an important contribution of this paper is to bring sequence-space perturbation methods to bear on optimal policy questions in heterogeneous-agent economies. We extend the fake-news algorithm of Auclert et al. (2021) to compute optimal policy and show that our timeless Ramsey approach is critical for the validity of sequence-space perturbations.

We show how to leverage both the primal and dual forms of the timeless Ramsey problem.[1] In the primal representation, we compute an extended set of sequence-space Jacobians and solve for the time paths of the multipliers that comprise a Ramsey plan. In the dual representation, we

[1] Throughout the paper, we say that a planning problem is in primal form when allocations or prices are explicit control variables for a planner, perhaps in addition to policy instruments. We say that a planning problem is in dual form when the only explicit control variables are the policy instruments. This terminology is consistent with standard use in related environments, e.g., Chari and Kehoe (1999) and Ljungqvist and Sargent (2018).

avoid having to compute the time paths of multipliers. However, approximating optimal policy in the dual is no longer possible in terms of sequence-space Jacobians and instead requires a second-order analysis. To that end, we introduce sequence-space Hessians as the natural, second-order generalization of sequence-space Jacobians.

Leveraging these methodological results, we conclude with a quantitative analysis of optimal monetary stabilization policy. Our approach allows us to compute transition dynamics under optimal policy—under discretion and with commitment—both non-linearly and to first order. We contrast optimal policy dynamics in HANK and RANK in response to demand shocks (Section 5.2), and productivity and cost-push shocks (Appendix F).

Related literature. Our paper contributes to multiple strands of the literature on optimal monetary policy, in particular recent work on optimal policy in HANK economies. Our continuous-time approach is most closely related to the work of Nuño and Thomas (2020), on which we build.[2,3] Nuño and Thomas (2020) study optimal monetary policy under commitment in a small open economy, in which short-term real interest rates and output are unaffected by policy.[4] Our paper studies optimal monetary policy in a closed economy that features the classic output-inflation tradeoff, which is central to the New Keynesian literature. Farhi and Werning (2016) study optimal monetary and macroprudential policies in a general heterogeneous-agent environment, highlighting the importance of labor wedges for optimal policies in the presence of nominal rigidities. Bhandari et al. (2021) introduce a small-noise expansion method to compute optimal monetary and fiscal policy in a HANK model with aggregate risk. Acharya et al. (2020) study optimal monetary policy in closed form in a HANK economy with constant absolute risk aversion (CARA) preferences and normally distributed shocks. Le Grand et al. (2021) study optimal monetary and fiscal policy keeping heterogeneity finite-dimensional by truncating idiosyncratic histories. González et al. (2021) characterize optimal monetary policy with heterogeneous firms. McKay and Wolf (2022) study optimal monetary policy with heterogeneous households in linear-quadratic environments.[5] Smirnov (2022) computes optimal monetary policy using a variational approach.

Our contribution relative to this body of work is fivefold. First, we provide the first analysis

[2] Nuño and Moll (2018) solve constrained-efficiency problems treating the cross-sectional distribution as a control.

[3] As emphasized by Werning (2011), continuous time Ramsey problems in New Keynesian economies are particularly tractable.

[4] Formally, the open economy setup in Nuño and Thomas (2020) immediately implies that both the Lagrange multipliers of the households' HJB equation and their optimality condition—which correspond to $\phi_t(a, z)$ and $\chi_t(a, z)$ in our paper, see equation (29)—are zero by construction. Hence, in their model, the planner would make the same savings decisions as households. Characterizing and computing these multipliers is a novel contribution of our paper.

[5] Several other papers study optimal monetary policy in environments with heterogeneity, typically relying on a second-order approximation to aggregate welfare. In particular, in two-agent New Keynesian environments, Bilbiie (2008, 2018) study optimal monetary policy without and with idiosyncratic risk, respectively; Cúrdia and Woodford (2016) study optimal monetary policy in a model with credit frictions; and Benigno et al. (2020) study optimal monetary policy at the zero lower bound.

of optimal policy under discretion in HANK economies, which shows that a utilitarian planner trades off aggregate stabilization against a novel redistribution motive. This redistribution motive is a new source of time inconsistency and exacerbates inflationary bias. Second, we jointly characterize three important dimensions of optimal monetary policy design: long-run policy, time consistency, and stabilization policy. Third, we introduce and analytically characterize the timeless penalties that resolve the time-0 problem of the standard Ramsey problem. A planner under discretion can implement the timeless Ramsey policy when confronted with the appropriate penalties. Relative to RANK, time consistent policy requires a novel distributional penalty. Fourth, the analytical targeting rules we derive for optimal monetary stabilization policy allow us to contrast policy prescriptions in HANK and RANK, which they nest as a special case. Finally, we extend the sequence-space approach to Ramsey problems, which allows us to compute optimal policy efficiently and fast.

We relate our results to the vast literature on monetary policy in RANK models and provide analytical insights into the departures of optimal policy from the RANK benchmark (Clarida et al., 1999; Woodford, 2003; Galí, 2015). At an abstract level, our approach is closest to Khan et al. (2003), who initially characterize standard and augmented (timeless) Ramsey plans and then use perturbation methods to characterize stabilization policy. Schmitt-Grohé and Uribe (2010) and Woodford (2010) systematically study and review optimal long-run policy and optimal stabilization policy in RANK economies. Our goal is to systematically revisit the canonical New Keynesian consensus on optimal monetary policy in the presence of household heterogeneity.

We formalize Woodford (1999)'s timeless perspective in our heterogeneous-agent setting by introducing a timeless penalty that resolves the standard Ramsey planner's time-0 problem (Kydland and Prescott, 1980).[6] Our characterization of the timeless penalty builds on the recursive multiplier approach of Marcet and Marimon (2019), which we extend to continuous-time heterogeneous-agent economies. Relative to RANK, we show that a time consistent implementation of monetary policy requires a novel distributional penalty and an inflation penalty augmented by distributional considerations. One contribution of our paper is to show that the distributional penalty solves a promise-keeping Kolmogorov forward equation.

Finally, we extend the sequence-space apparatus to Ramsey problems and welfare analysis, contributing to recent work on computational methods in heterogeneous-agent environments (Boppart et al., 2018; Auclert et al., 2021). We develop a sequence-space representation of timeless Ramsey plans, which we can compute non-linearly and using sequence-space perturbation methods. In particular, we extend the fake-news algorithm of Auclert et al. (2021) to compute Ramsey problems in both primal and dual forms. We also introduce and define sequence-space Hessians as the natural, second-order generalization of sequence-space Jacobians.

[6] See Woodford (2003, 2010) and Benigno and Woodford (2012) for expositions of the timeless perspective.

2 Model

Our baseline model is a one-asset heterogeneous-agent New Keynesian (HANK) model with wage rigidity (Auclert et al., 2018). It represents a minimal departure from a representative-agent New Keynesian (RANK) model (Clarida et al., 1999; Woodford, 2003; Galí, 2015).

Time is continuous and indexed by $t \in [0, \infty)$. There is no aggregate uncertainty and we focus on one-time, unanticipated shocks. Following much of the New Keynesian literature, we allow for demand, productivity, and cost-push shocks.

2.1 Households

The economy is populated by a unit mass of households whose lifetime utility is

$$V_0(\cdot) = \max \mathbb{E}_0 \int_0^\infty e^{-\int_0^t \rho_s ds} U_t(c_t, n_t) dt, \tag{1}$$

where U_t denotes the instantaneous utility flow from consumption c_t and labor n_t. Households discount at a potentially time-varying rate ρ_t, which represents a source of demand shocks. They can trade a single bond a_t and face a budget constraint

$$\dot{a}_t = r_t a_t + z_t w_t n_t + T_t(z_t) - c_t, \tag{2}$$

where r_t is the real interest rate and w_t the real wage rate. Beside financial and labor income, households may receive a lump-sum transfer $T_t(z_t)$ from the government, which will be zero in equilibrium, as described below. Finally, households face the borrowing constraint $a_t \geq \underline{a}$.

While there is no aggregate uncertainty, households face idiosyncratic earnings risk, captured by the exogenous Markov process z_t. Since we abstract from permanent heterogeneity, we can index individual households by their idiosyncratic state variables (a, z). We denote the mass of households in state (a, z) by $g_t(a, z)$, which we also refer to as the cross-sectional distribution.

2.2 Labor Market

As is standard in the New Keynesian sticky-wage literatures without heterogeneity (Erceg et al., 2000; Schmitt-Grohé and Uribe, 2005) and with heterogeneity (Auclert et al., 2018), labor unions determine work hours.[7] While Appendix A.1 details the union problem, we only summarize its relevant implications to study optimal monetary policy here. Labor is rationed, so all households supply the same hours,

$$n_t = N_t, \tag{3}$$

[7] It is possible to rederive our results in a model with price rigidity. The assumptions of sticky wages and symmetric labor rationing make our model more tractable. Moreover, firm profits are zero in equilibrium instead of counter-cyclical.

where N_t is aggregate labor. Nominal wages are sticky, and unions pay a quadratic Rotemberg (1982) adjustment cost to change wages. We assume this cost is passed to households as a utility cost, so that instantaneous flow utility in equation (1) takes the form

$$U_t(c_t, n_t) = u(c_t) - v(n_t) - \frac{\delta}{2}(\pi_t^w)^2, \tag{4}$$

where π_t^w denotes wage inflation and the parameter $\delta \geq 0$ modulates the degree of wage rigidity.

Unions choose wages to maximize stakeholder value—the private lifetime values of households. We show in Appendix A.1 that the union problem gives rise to the non-linear New Keynesian wage Phillips curve

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \underbrace{\frac{\epsilon_t}{\delta}}_{\text{NKPC slope}} \iint \Bigg(\underbrace{\frac{\epsilon_t - 1}{\epsilon_t} \overbrace{(1+\tau^L)}^{\text{Employment Subsidy}}}_{\text{Desired Markup}} z u'(c_t(a,z)) - \frac{v'(n_t)}{A_t} \Bigg) w_t n_t g_t(a,z)\, da\, dz, \tag{5}$$

where ϵ_t, the elasticity of substitution across unions, is potentially time-varying and a source of cost-push shocks. As is standard in the New Keynesian literature, we allow for a time-invariant employment subsidy τ^L to potentially offset unions' market power. This Phillips curve illustrates that labor wedges, which will play a key role in our welfare analysis, are key determinants of the dynamics of inflation. Formally, we define *individual inflation-relevant labor wedges* as $\tau_t(a,z) = (\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) z u'(c_t(a,z)) - \frac{v'(n_t)}{A_t}) w_t n_t$, and refer to $z u'(c_t(a,z)) - \frac{v'(n_t)}{A_t}$ as *individual* (welfare-relevant) *labor wedges*.[8] This definition allows us to rewrite the Phillips curve as

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta} \iint \tau_t(a,z) g_t(a,z)\, da\, dz, \tag{6}$$

which highlights that unions target an aggregate inflation-relevant labor wedge of zero.[9]

[8] Note that the inflation-relevant labor wedges are proportional to labor wedges when $\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) = 1$. In the limit as wages become flexible, $\delta \to 0$, there are no cost-push shocks, $\epsilon_t = \epsilon$, and we allow for the appropriate employment subsidy, the aggregate inflation- and welfare-relevant labor wedges coincide and are both zero.

[9] Equation (6) implies that an increase in the aggregate inflation-relevant labor wedge (on the RHS) leads to an increase in the rate of change of inflation (on the LHS). Since the Phillips curve is a forward-looking equation with a terminal condition, an increase in the rate of change of inflation requires a fall in the actual level of inflation. This is consistent with the interpretation of negative aggregate labor wedges indicating a recession, which generates deflationary pressure. The aggregate inflation-relevant labor wedge can rise either if individual inflation-relevant labor wedges $\tau_t(a,z)$ increase or mass $g_t(a,z)$ shifts to states with high $\tau_t(a,z)$.

2.3 Final Good Producer

A representative firm produces the final consumption good using labor,

$$Y_t = A_t N_t, \tag{7}$$

where total factor productivity (TFP) A_t is potentially time-varying and a source of exogenous productivity shocks. Under perfect competition and flexible prices, profits from production are zero and the marginal cost of labor is equal to its marginal product, with

$$w_t = A_t, \tag{8}$$

so the real wage w_t is equal to the marginal rate of transformation (MRT) A_t.

2.4 Government

The role of fiscal policy is deliberately minimal. There is no government spending and no debt, with bonds in zero net supply. The fiscal authority pays an employment subsidy $\tau^L z_t w_t n_t$ to households with labor productivity z_t. To balance the budget, it raises a lump-sum tax also proportional to labor productivity as well as aggregate labor income. Households therefore receive a net fiscal rebate of $T_t(z_t) = \tau^L z_t w_t n_t - \tau^L z_t w_t N_t = 0$. Our focus is instead on the monetary authority, which optimally sets the path of nominal interest rates $\{i_t\}_{t\geq 0}$. This is the only policy instrument of the planner. A Fisher relation holds, with

$$r_t = i_t - \pi_t, \tag{9}$$

where π_t is consumer price inflation. Finally, we can relate price inflation to wage inflation by differentiating equation (8), which yields

$$\pi_t = \pi_t^w - \frac{\dot{A}_t}{A_t}. \tag{10}$$

2.5 Equilibrium and Implementability

Definition 1. (**Competitive Equilibrium**) *Given an initial distribution over household bond holdings and idiosyncratic labor productivities, $g_0(a, z)$, and given predetermined sequences of monetary policy $\{i_t\}$ and shocks $\{A_t, \rho_t, \epsilon_t\}$, an equilibrium is defined as paths for prices $\{\pi_t^w, \pi_t, w_t, r_t\}$, aggregates $\{Y_t, N_t, C_t, B_t\}$, individual allocation rules $\{c_t(a, z)\}$, and the distribution $\{g_t(a, z)\}$ such that households optimize, unions* *optimize, labor is rationed, firms optimize, and markets for goods and bonds clear, that is,*

$$Y_t = C_t = \iint c_t(a,z)\, g_t(a,z)\, da\, dz \tag{11}$$

$$0 = B_t = \iint a\, g_t(a,z)\, da\, dz. \tag{12}$$

The Ramsey problems we study in Sections 3 and 4 take a primal approach: the planner chooses among those competitive equilibria that are implementable by policy. We formally derive these implementability conditions in Lemma 12 in Appendix A and show there that they comprise five equations—three at the individual level and two at the aggregate level.

At the individual level, the planner must respect individuals' consumption-savings decisions. These are encoded in the household's standard first-order condition for consumption, which equates marginal utility of consumption with the private marginal value of wealth,

$$u'(c_t(a,z)) = \partial_a V_t(a,z). \tag{13}$$

Because equation (13) features the private lifetime value $V_t(a,z)$, the planner must also respect its evolution over time. A standard Bellman equation—or Hamilton-Jacobi-Bellman (HJB) equation in continuous time—relates current lifetime value $V_t(a,z)$ to flow utility and continuation value,

$$\rho_t V_t(a,z) = \underbrace{u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2}_{\text{Flow Utility}} + \underbrace{\partial_t V_t(a,z) + \mathcal{A}_t V_t(a,z)}_{\text{Continuation Value}}. \tag{14}$$

The continuation value from state (a,z) at time t is $\mathbb{E}_t\left[\frac{dV_t(a,z)}{dt}\right] = \partial_t V_t(a,z) + \mathcal{A}_t V_t(a,z)$, where $\mathcal{A}_t$ denotes the infinitesimal generator of the process (a_t, z_t), formally defined in equation (60) in Appendix A.4.[10] Finally, the planner internalizes that a change in policy affects the evolution of the cross-sectional household distribution, characterized in continuous time by the Kolmogorov forward (KF) equation

$$\partial_t g_t(a,z) = \mathcal{A}_t^* g_t(a,z), \tag{15}$$

where $\mathcal{A}_t^*$ denotes the adjoint of $\mathcal{A}_t$.[11] The KF equation tracks the movement of households across individual states (a,z) over time. The relationship between the generator $\mathcal{A}_t$ and its adjoint $\mathcal{A}_t^*$ connects the Bellman equation (14) and the KF equation (15): Under a law of large numbers, a household's rational expectations over future transitions across states must be consistent with the actual evolution of the cross-sectional distribution.

[10] Using households' Bellman equations and consumption-savings optimality conditions as implementability conditions instead of consumption Euler equations is critical to derive analytical insights.

[11] The adjoint of an operator can be seen as a generalization of the transpose of a matrix.

At the aggregate level, the implementability conditions comprise the aggregate resource constraint that combines the goods market clearing condition (11) with the production function (7),

$$\iint c_t(a,z)\, g_t(a,z)\, da\, dz = A_t N_t, \tag{16}$$

as well as the New Keynesian wage Phillips curve (6).

Comparison benchmarks. Our HANK model nests two useful benchmarks. First, we relate our results to the RANK limit of our model (Appendix E). Second, we define the flexible-wage benchmark as the limit of our economy as $\delta \to 0$, analogous to the flexible-price limit in the canonical New Keynesian analysis (Appendix A.3). We refer to natural output as the output that obtains in the flexible-wage limit, denoted $\tilde{Y}_t$.

2.6 Sources of Suboptimality

To conclude the description of the model, we discuss its four sources of suboptimality.[12]

1. First, monopolistic competition drives a wedge between the real wage, which is equal to the marginal rate of transformation (MRT), A_t, and households' average marginal rate of substitution (MRS) between consumption and labor. The appropriate employment subsidy may offset this wedge in steady state.

2. Second, nominal wage rigidity implies that the economy's average MRS can converge only gradually to the MRT in response to shocks. Moreover, wage adjustment costs represent a direct deadweight loss (utility cost to households).

3. Third, our model also features labor rationing. In the absence of aggregate shocks and with the appropriate employment subsidy, an appropriate notion of average MRS is equal to the MRT in our economy. However, individual MRS are not equalized across households because all households are required to work the same hours.

4. Finally, and most importantly, there are incomplete markets for risk: Noncontingent bonds are the only financial asset in this economy and households face a borrowing constraint, jointly restricting their ability to self-insure against idiosyncratic earnings risk. Both forms of incompleteness imply that households' marginal rates of substitution are not equalized across periods and states.

[12] This subsection is meant to parallel Section 4 of Khan et al. (2003) and Chapter 4.2 of Galí (2015), which discuss sources of suboptimality in RANK economies.

The first two sources of inefficiency exactly mirror those in the standard New Keynesian model. Labor rationing and incomplete markets, on the other hand, are unique to the heterogeneous-agent environment. While labor rationing is not critical for the insights of this paper, and could be eliminated at the cost of complicating the labor market block, the presence of incomplete markets is central to our analysis. Due to these two inefficiencies, the flexible-wage allocation is no longer first-best in a HANK economy even with the appropriate employment subsidy.

3 Optimal Monetary Policy under Discretion

We structure our analysis of optimal monetary policy to parallel that of Clarida et al. (1999), starting with policy under discretion in Section 3 and studying policy with commitment in Section 4. Throughout the paper, we adopt an equal-weighted utilitarian welfare criterion. While this assumption is not innocuous, it is a natural starting point—see Dávila and Schaab (2022) for a systematic study of welfare criteria in general dynamic stochastic environments.[13]

Ramsey problem with finite commitment horizon. Under discretion, a planner has control over policy in the present and takes future policy—under the control of a future planner—as well as agents' expectations as given. In discrete time, it is straightforward to associate the present with period t and the future with periods $t+1$ and onwards (Clarida et al., 1999). To stay as close as possible to this notion of policy under discretion in continuous time, we introduce a *Ramsey problem with finite commitment horizon*.

Formally, we consider a planner who exercises control (and has commitment) over policy over some finite time horizon—the analog of the time interval $[t, t+1)$ in discrete time. At the transition time, which occurs at the transition rate ψ, the present planner is replaced by another who sets policy going forward until she herself is again replaced. Planners do not honor promises made by previous planners. We denote the times at which planners transition by $\{\tau_n\}_{n=0}^{\infty}$, with $\tau_0 = 0$ the starting time of the first planner.[14]

Definition 2. (**Ramsey Problem with Finite Commitment Horizon**) *A Ramsey planner with finite commitment horizon $[0, \tau_1)$ chooses allocations, prices, and policy*

$$\boldsymbol{X} = \{c_t(a,z), V_t(a,z), g_t(a,z), N_t, \pi_t^w, i_t\}_{t=0}^{\tau_1}$$

[13] In ongoing work, we study optimal monetary policy and central bank mandates under alternative welfare criteria (Dávila and Schaab, 2023).

[14] Formally, our infinitesimal discretion approach merges insights from Marcet and Marimon (2019) with the continuous-time results of Harris and Laibson (2013). Schaumburg and Tambalotti (2007) study a similar planning problem with finite commitment horizon in a RANK model in discrete time. See Appendix C for details.

as well as multipliers

$$\boldsymbol{M} = \{\phi_t(a,z), \chi_t(a,z), \lambda_t(a,z), \mu_t, \theta_t\}_{t=0}^{\tau_1}$$

to maximize social welfare subject to implementability conditions (6, 13 – 16), taking as given the initial cross-sectional distribution $g_0(a,z)$ as well as future policy. That is,

$$\mathcal{W}_0(g_0) = \min_{\boldsymbol{M}} \max_{\boldsymbol{X}} \mathbb{E}_0 \Big[L(0, \tau_1, g_0) + e^{-\int_0^{\tau_1} \rho_s ds} \mathcal{W}_{\tau_1}(g_{\tau_1}) \Big] \tag{17}$$

where the expectation $\mathbb{E}_0$ is over the transition time τ_1, $\mathbb{E}_0\big[e^{-\int_0^{\tau_1}\rho_s ds}\mathcal{W}_{\tau_1}(g_{\tau_1})\big]$ denotes the expected discounted continuation value, and $L(t_1, t_2, g_{t_1})$ is the planner's Lagrangian over the horizon $[t_1, t_2)$, given an initial cross-sectional distribution $g_{t_1}(a,z)$:

$$\begin{aligned} L(t_1, t_2, g_{t_1}) = \int_{t_1}^{t_2} e^{-\int_{t_1}^{t} \rho_s ds} \Bigg\{ \iint \Bigg\{ & U_t(a,z)\, g_t(a,z) \\ & + \phi_t(a,z) \Big[-\rho_t V_t(a,z) + U_t(a,z) + \partial_t V_t(a,z) + \mathcal{A}_t V_t(a,z) \Big] \\ & + \chi_t(a,z) \Big[u'(c_t(a,z)) - \partial_a V_t(a,z) \Big] \\ & + \lambda_t(a,z) \Big[-\partial_t g_t(a,z) + \mathcal{A}_t^* g_t(a,z) \Big] \Bigg\} da\, dz \\ & + \mu_t \Big[\iint (c_t(a,z) - A_t z N_t) g_t(a,z)\, da\, dz \Big] \\ & + \theta_t \Big[-\partial_t \pi_t^w + \rho_t \pi_t^w + \frac{\epsilon_t}{\delta} \iint \tau_t(a,z) g_t(a,z)\, da\, dz \Big] \Bigg\} dt, \end{aligned} \tag{18}$$

where $U_t(a,z) = u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2$. The operators $\mathcal{A}_t$ and $\mathcal{A}_t^*$ are defined in Appendix A.4.

In the remainder of the paper, we focus on two limits of this finite-horizon Ramsey problem: First, as $\psi \to 0$, planners never transition. In fact, the first planner stays in power forever. The resulting Ramsey problem is thus simply the standard Ramsey problem with an infinite commitment horizon as we show in Section 4.1. Second, as $\psi \to \infty$, planners transition increasingly frequently and their commitment horizon becomes vanishingly small. This is the limit we associate with *policy under discretion* in continuous time.[15]

[15] Formally, for a given ψ, we study the Markov perfect equilibrium of the game played by a sequence of Ramsey planners with finite commitment horizon. It comprises i) paths for prices, π_t^w, aggregates, N_t, individual consumption allocations and value functions, $c_t(a,z)$ and $V_t(a,z)$, as well as cross-sectional distributions, $g_t(a,z)$, that satisfy the competitive equilibrium conditions (6, 13 – 16) given paths for policy, i_t, and shocks, $(A_t, \rho_t, \epsilon_t)$, as well as an initial distribution $g_0(a,z)$; ii) a path of interest rate policy i_t; and iii) a sequence of multiplier functions, $\phi_t(a,z)$, $\chi_t(a,z)$,

The implementability conditions encoded in the Lagrangian (18) include two forward-looking constraints: the individual Bellman equations and the Phillips curve, respectively associated with the multipliers $\phi_t(a,z)$ and θ_t. In Clarida et al. (1999), lack of commitment implies the planner takes as given next-period inflation expectations. In our continuous-time formulation, the planner similarly takes as given expectations about inflation and value assignments beyond the current commitment horizon. Formally, this is encoded in the terminal conditions π_{τ_1} and $V_{\tau_1}(a,z)$ that the planner faces, and which are themselves part of the solution of Markov perfect equilibrium as in discrete time.[16]

Optimal monetary policy under discretion: optimality conditions and interpretation. We now summarize the necessary first-order conditions that characterize optimal monetary policy under discretion, i.e., in the limit of the finite-horizon Ramsey problem (17) as $\psi \to \infty$.

Proposition 1. (**Policy under Discretion: Optimality Conditions**) *The necessary first-order conditions that characterize optimal monetary policy under discretion are given by*

$$\rho_t \lambda_t(a,z) = U_t(a,z) + \mu_t(c_t(a,z) - A_t z N_t) + \partial_t \lambda_t(a,z) + \mathcal{A}_t \lambda_t(a,z) \tag{19}$$

$$0 = u'(c_t(a,z)) - \partial_a \lambda_t(a,z) + \mu_t - \tilde{\chi}_t(a,z) \tag{20}$$

$$0 = \iint z \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{z}\, \zeta_t^{HTM}\, g_t(\underline{a},\underline{z})\, da\, dz - \mu_t - \frac{v'(N_t)}{A_t} \tag{21}$$

$$0 = \iint a \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{a}\, \zeta_t^{HTM}\, g_t(\underline{a},\underline{z})\, da\, dz \tag{22}$$

where

$$\zeta_t^{HTM} \equiv u'(c_t(\underline{a},\underline{z})) - \partial_a \lambda_t(\underline{a},\underline{z}) + \mu_t \quad \text{and} \quad \tilde{\chi}_t(a,z) \equiv -u''(c_t(a,z)) \frac{\chi_t(a,z)}{g_t(a,z)}.$$

In the limit as $\psi \to \infty$, the paths of the multipliers on forward-looking implementability conditions converge to $\theta_t \to 0$ and $\phi_t(a,z) \to 0$ for all t and (a,z).

Equations (19) through (22) correspond to the first-order conditions for the cross-sectional distribution $g_t(a,z)$, individual consumption $c_t(a,z)$, aggregate activity N_t, and the nominal interest rate i_t,

$\lambda_t(a,z)$, μ_t, and θ_t that solve (17). Policy under discretion corresponds to the limit of this equilibrium as $\psi \to \infty$ and planners have vanishingly small commitment horizons.

[16] In this problem, there are two direct linkages between the present finite-horizon Ramsey problem and future policy. The first is encoded in the continuation value $\mathcal{W}_{\tau_1}(g_{\tau_1})$: the present Ramsey planner internalizes that choosing policy today affects the evolution of the cross-sectional distribution and, therefore, the initial condition $g_{\tau_1}(a,z)$ of the future planner at the time of transition. Second, taking future policy as given implies that the present planner faces terminal conditions for each forward-looking constraint. Concretely, the planner takes as given inflation π_{τ_1} and values $V_{\tau_1}(a,z)$ at the time of transition. This is analogous to the setup in Clarida et al. (1999), where the present planner takes as given inflation expectations.

respectively. Judiciously combining these conditions allows us to characterize the properties of optimal monetary policy under discretion.

To that end, we start by providing an economic interpretation of the three non-zero multipliers $\chi_t(a,z)$, $\lambda(a,z)$, and μ_t. First, the multiplier $\chi_t(a,z)$ corresponds to the social shadow value of relaxing households' consumption-savings decisions. When $\chi_t(a,z) > (<) \; 0$, the planner perceives that households in state (a,z) save (consume) too much relative to an environment in which the planner could perfectly manage consumption-savings decisions. The multiplier $\chi_t(a,z)$ acts as the shadow penalty that ensures the planner respects private consumption-savings decisions. Second, the multiplier $\lambda_t(a,z)$ corresponds to the social shadow value of increasing the mass of households in state (a,z). As we show below, this multiplier represents the social lifetime value of a household in state (a,z). Third, the multiplier μ_t corresponds to the social shadow value of increasing aggregate excess demand. When $\mu_t > (<) \; 0$, the planner perceives that increasing (reducing) aggregate demand or reducing (increasing) aggregate supply is socially beneficial.[17]

After introducing the non-zero multipliers, we interpret the four optimality conditions (19) through (22). First, equation (19) implies that $\lambda_t(a,z)$ defines the social lifetime value of a household.[18] The difference between private lifetime value (14) and social lifetime value under discretion (19) is given by the term $\mu_t(c_t(a,z) - A_t z N_t)$, which captures the contribution of a household in state (a,z) to aggregate excess demand. Intuitively, households for whom $c_t(a,z) > A_t z N_t$ put positive pressure on aggregate excess demand since their contribution to aggregate demand, $c_t(a,z)$, is higher than their contribution to aggregate supply, $w_t z N_t$, which is socially desirable (undesirable) when $\mu_t > (<) \; 0$. Equation (19) allows us to characterize the social marginal value of wealth—a key input for the remaining optimality conditions—as

$$\partial_a \lambda_t(a,z) = \partial_a V_t(a,z) + \mathcal{M}_t(a,z)\mu_t, \tag{23}$$

where $\mathcal{M}_t(a,z)$ denotes an operator, introduced in Appendix A.4, that acts on the path of multipliers μ_t. The difference between the private and the social marginal value of wealth, $\mathcal{M}_t(a,z)\mu_t$, can be interpreted as the present discounted value of the contribution of future consumption to aggregate excess demand induced by an increase in the household's wealth at time t.[19]

Second, equation (20) has the interpretation of a social consumption-savings optimality con-

[17] The interpretation of ξ_t^{HTM} is the social marginal value of giving a dollar of (unearned) income to every hand-to-mouth household, whose mass is $g_t(\underline{a},\underline{z})$. For our proofs, we assume that $z_t \in \{\underline{z}, \overline{z}\}$ follows a two-state Markov chain, so only households at the borrowing constraint $\underline{a}$ and with the low earnings realization $\underline{z}$ are hand-to-mouth (Achdou et al., 2022). For the perturbations that feature ξ_t^{HTM}, we are holding fixed consumption for all households when, e.g., perturbing i_t, except for the hand-to-mouth households: The planner must respect the borrowing constraint and cannot freely choose the consumption of households at the borrowing constraint. We therefore consider perturbations where, only for the hand-to-mouth household, a change in income leads to a change in consumption.

[18] The multiplier $\lambda_t(a,z)$ takes the form of an HJB equation and can alternatively be written as $\rho\lambda_t = U_t + \mu_t(c_t - w_t z N_t) + \mathbb{E}_t\left[\frac{d\lambda_t}{dt}\right]$, suppressing the dependence on (a,z).

[19] $\mathcal{M}_t(a,z)$ is positive and bounded between 0 and 1 under mild regularity conditions. See Appendix A.4.

dition. Like households, the planner trades off the direct benefit of increasing consumption, $u'(c_t(a,z))$, against the marginal value of having higher future assets from savings, which the planner values at $\partial_a \lambda_t(a,z)$. Moreover, increasing consumption increases aggregate excess demand, which is socially desirable (undesirable) when $\mu_t > (<)\ 0$. Since the planner must also respect private consumption-savings decisions, $\tilde{\chi}_t(a,z) > (<)\ 0$ acts as an additional social shadow cost (benefit) of consumption that ensures that the individual consumption-savings optimality condition is satisfied.

Third, equation (21) has the interpretation of an aggregate activity condition, and represents the planner's valuation of an increase in hours worked by all households, which has three components.[20] First, household wealth increases in proportion to the effective wage $zw_t = zA_t$. The planner values this effect using the social marginal value of wealth for unconstrained households, $\partial_a \lambda_t(a,z)$, and the social marginal value of consumption for constrained households, $\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t^{HTM}(\underline{a},\underline{z})$.[21] Second, aggregate supply increases by $\iint zA_t g_t(a,z)\, da\, dz = A_t$, which the planner values at the shadow value of aggregate excess demand, μ_t. Third, the planner accounts for households' direct disutility from working more, $\iint v'(N_t) g_t(a,z)\, da\, dz = v'(N_t)$. When choosing optimal aggregate economic activity, the planner trades off these three forces.

Finally, equation (22) represents the planner's valuation of an increase in the nominal interest rate. In this environment, an increase in the interest rate redistributes dollars across households in proportion to their bond holdings a. The planner values such redistribution in dollars according to $\partial_a \lambda_t(a,z)$ for unconstrained households and $\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t^{HTM}$ for constrained households. This term captures the distributive pecuniary effect of a change in interest rates, which is central to the determination of optimal monetary policy, as we show next.[22]

Targeting rule and inflationary bias. Combining the optimality conditions just described allows us to characterize a *targeting rule* for optimal monetary policy under discretion in Proposition 2. This targeting rule illustrates the forces that optimal monetary policy under discretion balances in our HANK model and facilitates the comparison of our results to the canonical analysis of monetary policy in RANK, which we show to be nested as a special case by our targeting rule. We

[20] Since the planner must respect how the union allocates labor, the planner can only consider perturbations that change hours worked for all households symmetrically.

[21] The social value of increasing the consumption of a household at the borrowing constraint corresponds to

$$\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t^{\mathrm{HTM}} = u'(c_t(\underline{a},\underline{z})) + \mu_t. \tag{24}$$

Intuitively, the planner internalizes that a marginal change in the wealth of a household at the borrowing constraint leads to a one-for-one change in consumption, whose social value is given by the sum of the direct utility benefit, $u'(c_t(\underline{a},\underline{z}))$, and the impact on aggregate excess demand, μ_t.

[22] We use the terminology distributive pecuniary effects as in Dávila and Korinek (2018). That paper shows that distributive pecuniary effects are characterized by i) changes in net asset positions, here a, and ii) differences in valuation, here $\partial_a \lambda_t(a,z)$. As shown in that paper, if the planner valued a dollar across households identically, market clearing would imply that distributive pecuniary effects are zero, so $\iint a \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{a}\, \xi_t^{\mathrm{HTM}}\, g_t(\underline{a},\underline{z})\, da\, dz = 0$.

present a constructive proof of Proposition 2 in Appendix C.2.

Proposition 2. (**Targeting Rule under Discretion**) *Optimal monetary policy under discretion is characterized by the non-linear targeting rule*

$$\underbrace{\iint \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) g_t(a,z)\, da\, dz}_{\text{Aggregate Labor Wedge}} = \Omega_t^{\mathrm{D}} \underbrace{\iint a u'(c_t(a,z)) g_t(a,z)\, da\, dz}_{\text{Distributive Pecuniary Effect}}, \tag{25}$$

where Ω_t^{D}*, given by*

$$\Omega_t^{\mathrm{D}} = \frac{\iint z(1-\mathcal{M}_t(a,z)) g_t(a,z)\, da\, dz - (1-\mathcal{M}_t(\underline{a},\underline{z}))\, \underline{z}\, g_t(\underline{a},\underline{z})}{\iint a(1-\mathcal{M}_t(a,z)) g_t(a,z)\, da\, dz - (1-\mathcal{M}_t(\underline{a},\underline{z}))\, \underline{a}\, g_t(\underline{a},\underline{z})},$$

is positive under mild regularity conditions.

This non-linear targeting rule shows that, under discretion, the utilitarian planner trades off aggregate stabilization against a novel redistribution motive. The LHS of equation (25) is the aggregate labor wedge and represents the aggregate stabilization motive of the planner, while the RHS is the distributive pecuniary effect of interest rate changes. Crucially, the marginal utility of consumption falls with household wealth, so that

$$\iint a u'(c_t(a,z)) g_t(a,z)\, da\, dz = \mathbb{C}ov_{g_t(a,z)}\Big(a, u'(c_t(a,z))\Big) < 0,$$

where $\mathbb{C}ov_{g_t(a,z)}(a, u'(c_t(a,z)))$ is the cross-sectional covariance between wealth and marginal utility.[23]

Therefore, optimal monetary policy under discretion targets a negative aggregate labor wedge, which is associated with an overheated economy. To illustrate, consider a level of interest rates at which the aggregate labor wedge is zero and policy attains aggregate stabilization. The negative RHS of (25) implies that, relative to the policy stance under consideration, the planner finds it valuable to lower the real interest rate in order to redistribute towards indebted, high marginal utility households. To lower the real rate, the planner lowers the nominal policy rate, which results in a negative aggregate labor wedge, i.e., an overheated economy. Proposition 2 thus offers a novel perspective on optimal monetary policy under discretion in a heterogeneous-agent environment.[24]

[23] While the positive HANK literature has concluded that marginal propensities to consume (MPC) are central for monetary policy transmission, equation (22) highlights that marginal utilities are instead the key direct determinant of the targeting rules for optimal monetary policy. This is the case under discretion, but also with commitment—see Section 4.

[24] In the RANK limit of our economy, the planner's motive to redistribute via distributive pecuniary interest rate effects vanishes. Formally, the RHS of equation (25) goes to 0 in that limit and, as a result, optimal monetary policy under discretion focuses solely on aggregate stabilization, targeting an aggregate labor wedge of 0.

When assuming isoelastic (CRRA) preferences, with $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$ and $v(n) = \frac{1}{1+\eta} n^{1+\eta}$, the targeting rule under discretion takes the form

$$Y_t = \tilde{Y}_t \times \underbrace{\left(\frac{\epsilon_t}{\epsilon_t - 1} \frac{1}{1 + \tau^L} \right)^{\frac{1}{\gamma+\eta}}}_{\text{Desired Markup: } \geq 1} \times \underbrace{\left(1 - \Omega_t^{\mathrm{D}} \frac{\iint a u'(c_t(a,z)) g_t(a,z) \, da \, dz}{\iint z u'(c_t(a,z)) g_t(a,z) \, da \, dz} \right)^{\frac{1}{\gamma+\eta}}}_{\text{Desired Redistribution: } > 1} \tag{26}$$

where $\tilde{Y}_t$ denotes natural output as defined in equation (59). While this output gap targeting rule is more closely connected to the canonical results on optimal monetary policy in RANK, equations (25) and (26) have the same content.

Equation (26) shows that, under discretion, monetary policy targets output to be equal to natural output, i.e., to close the output gap, up to two wedges. The first wedge is the familiar one deriving from monopolistic competition and unions' desired markups, due to which employment may be inefficiently low. Whenever the employment subsidy τ^L is not sufficiently large, this wedge is positive, motivating the planner to raise output above potential to raise employment.

In HANK, a second redistribution wedge emerges, since marginal utility of consumption falls with wealth, i.e., $\mathbb{C}ov_{g_t(a,z)}(a, u'(c_t(a,z))) < 0$. This wedge is therefore strictly positive, encouraging the utilitarian planner under discretion to overheat the economy even further.

An important conclusion of the canonical monetary policy analysis in RANK is that there are no gains from commitment when the planner sets the correct steady state employment subsidy and there are no cost-push shocks. Indeed, in the RANK limit of our economy, $\Omega_t^{\mathrm{D}} \to 0$ and the redistribution wedge vanishes. And when $\frac{\epsilon_t}{\epsilon_t - 1} \frac{1}{1+\tau^L} = 1$, equation (26) collapses to $Y_t = \tilde{Y}_t$: In that case, in RANK, monetary policy under discretion closes the output gap, which also closes the inflation gap and Divine Coincidence obtains even without commitment. In HANK, this is no longer the case as a planner under discretion always has an incentive to overheat the economy due to distributional considerations.

In equilibrium, agents anticipate the planner's incentive to raise output above natural output, rendering the planner's attempt to stimulate the economy futile. Instead, inflation ensues. Proposition 3 shows that a Markov perfect stationary equilibrium features inflationary bias, now exacerbated by the novel redistribution motive.

Proposition 3. (**Inflationary Bias**) *The Markov perfect stationary equilibrium with optimal monetary policy under discretion features inflationary bias, given by*

$$\pi_{ss}^w = \frac{\epsilon}{\delta} A_{ss} N_{ss} \Bigg[\underbrace{\left(1 - \frac{\epsilon - 1}{\epsilon} (1 + \tau^L) \right) \Lambda_{ss}}_{\textit{Markup: } \geq 0} - \Omega_{ss}^D \underbrace{\mathbb{C}ov_{g_{ss}(a,z)} \Big(a, u'(c_{ss}(a,z)) \Big)}_{\textit{Redistribution: } < 0} \Bigg]. \tag{27}$$

Consistent with our discussion of targeting rules under discretion, inflationary bias emerges from two sources: inefficiently low employment due to markups, as in the RANK limit, and redistribution. Quantitatively, the contribution of the novel redistribution motive is over 4 times larger than that of markups in our calibration exercise—see Figure 1 in Section 4.3.

Our stylized model features a single distributive pecuniary effect associated with adjusting interest rates. In richer models, optimal policy would account for all pecuniary effects and inflationary bias would be determined by the covariance between marginal utility and the aggregate of those effects. While our approach and the logic behind our results apply more generally, the exact quantitative conclusions—including whether policy under discretion features an inflationary or deflationary bias—may not.

Practical implications of optimal monetary policy under discretion. Summing up, our analysis of optimal monetary policy under discretion yields three main takeaways. First, a utilitarian planner has an incentive to run an overheated economy. This occurs because the planner values redistribution toward indebted, high marginal utility households via lower interest rates. The planner trades off this novel redistribution motive against aggregate stabilization. Second, the economy features inflationary bias in the sense of Barro and Gordon (1983). The standard motive to stimulate the economy due to markups is exacerbated by a novel desire to redistribute towards indebted households. When agents anticipate these incentives in equilibrium, both result in elevated inflation. Quantitatively, the redistribution motive is the dominant source of inflationary bias. Third, the markup-correcting employment subsidy, $\frac{\epsilon-1}{\epsilon}(1+\tau^L)=1$, that eliminates inflationary bias in RANK is no longer sufficient to address inflationary bias in HANK.[25]

4 Optimal Monetary with Policy with Commitment

We have shown in Section 3 that the desire to redistribute exacerbates inflationary bias when a utilitarian planner sets optimal policy under discretion. In this section, we characterize optimal monetary policy under commitment. We proceed in three steps.

First, in Section 4.1, we introduce the standard Ramsey problem and characterize the associated Ramsey plan and stationary Ramsey plan. In particular, we show in Section 4.2 that the optimal stationary equilibrium under commitment features zero inflation, eliminating the inflationary bias of policy under discretion in the long run.

Second, in Section 4.3, we show that while the full-commitment standard Ramsey problem eliminates inflationary bias in the long run, it still suffers from inflationary bias in the short run. This is due to a "time-0 problem" (Kydland and Prescott, 1980) associated with two dimensions

[25] In principle, it is possible to set a sufficiently large employment subsidy τ_L so that inflationary bias is zero.

of time inconsistency. In order to find a "timeless" planning solution, we extend the approach of Marcet and Marimon (2019) to our setting (i.e., continuous-time heterogeneous-agent economies) by introducing timeless penalties for each forward-looking implementability condition. We then define a timeless Ramsey problem, which augments the standard Ramsey problem with the timeless penalties, and prove that it no longer suffers from a time-0 problem—there is no incentive to deviate from the stationary Ramsey plan in the absence of shocks. Hence, the timeless Ramsey problem resolves inflationary bias in both the short run and the long run. In Sections 4.4 and 4.5, we study the determinants of the timeless penalty and contrast implementations of optimal policy based on penalties and targets.

Finally, in Section 4.6, we characterize optimal stabilization policy under the timeless Ramsey problem, which allows us to separate the pure stabilization motive from the time-0 problem.

Each of these three steps isolates one important dimension of optimal monetary policy design. First, characterizing the stationary Ramsey plan allows us to solve for optimal long-run policy, with which the planner addresses distortions in a stationary equilibrium. Second, characterizing the timeless Ramsey plan and the timeless penalties that support it allows us to isolate the planner's incentives to deviate from the stationary Ramsey plan in the short run due to the time-0 problem. Finally, by characterizing optimal stabilization policy with the appropriate timeless penalties, we isolate the planner's pure stabilization motive, no longer confounded by long-run distortions and time inconsistency considerations.

4.1 Standard Ramsey Problem and Ramsey Plan

The standard Ramsey problem corresponds to the limit of problem (17) as $\psi \to 0$, i.e., as the commitment horizon becomes infinite. We state the full problem in Appendix B.1 for convenience.

Definition 3. (**Standard Primal Ramsey Problem / Ramsey Plan**)

a) *The standard primal Ramsey problem solves*

$$\min_{\{\phi_t(a,z),\chi_t(a,z),\lambda_t(a,z),\mu_t,\theta_t\}} \max_{\{c_t(a,z),V_t(a,z),g_t(a,z),N_t,\pi_t^w,i_t\}} L^{\mathrm{SP}}(g_0) \tag{28}$$

where $L^{\mathrm{SP}}(g_0)$ denotes the standard primal Lagrangian, given an initial distribution of bond holdings and idiosyncratic labor productivity $g_0(a,z)$:

$$L^{\mathrm{SP}}(g_0) = \lim_{T\to\infty} L(0,T,g_0). \tag{29}$$

b) *A Ramsey plan corresponds to the solution of this problem and comprises i) paths for prices, π_t^w, aggregates, N_t, individual consumption allocations and value functions, $c_t(a,z)$ and $V_t(a,z)$, and* *the cross-sectional distribution, $g_t(a,z)$, that satisfy the implementability conditions given paths for interest rates, i_t, and shocks, $(A_t, \rho_t, \epsilon_t)$, as well as an initial distribution, $g_0(a,z)$; ii) a path of interest rate policy i_t; and iii) paths for the multiplier functions, $\phi_t(a,z)$, $\chi_t(a,z)$, $\lambda_t(a,z)$, μ_t, and θ_t that solve (28).*

Proposition 4 summarizes the optimality conditions that characterize the standard Ramsey plan. Our derivation relies on a variational approach, formally developed in Appendix B.[26]

Proposition 4. (**Standard Primal Ramsey Problem: Optimality Conditions**) *The optimality conditions for the standard primal Ramsey problem are given by*

$$\partial_t \phi_t(a,z) = -\mathcal{A}_t^* \phi_t(a,z) + \partial_a \chi_t(a,z) \tag{30}$$

$$\rho_t \lambda_t(a,z) = U_t(a,z) + \mu_t(c_t(a,z) - A_t z N_t) + \theta_t \frac{\epsilon_t}{\delta} \tau_t(a,z) + \partial_t \lambda_t(a,z) + \mathcal{A}_t \lambda_t(a,z) \tag{31}$$

$$0 = u'(c_t(a,z)) - \partial_a \lambda_t(a,z) + \mu_t + \theta_t \frac{\epsilon_t}{\delta} \frac{d\tau_t(a,z)}{dc_t(a,z)} - \tilde{\chi}_t(a,z) \tag{32}$$

$$\begin{aligned} 0 = & \iint z \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{z} \xi_t^{HTM} g_t(\underline{a}, \underline{z})\, da\, dz - \mu_t - \frac{v'(N_t)}{A_t} \\ & + \iint \phi_t(a,z) \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) da\, dz + \theta_t \frac{\epsilon_t}{\delta} \iint \frac{1}{A_t} \frac{d\tau_t(a,z)}{dN_t} g_t(a,z)\, da\, dz \end{aligned} \tag{33}$$

$$\dot{\theta}_t = \delta \pi_t^w \left(1 + \iint \phi_t(a,z)\, da\, dz \right) \tag{34}$$

$$0 = \iint \left(a \partial_a \lambda_t(a,z) g_t(a,z) + a \phi_t(a,z) \partial_a V_t(a,z) \right) da\, dz + \underline{a} \xi_t^{HTM} g_t(\underline{a}, \underline{z}) da\, dz \tag{35}$$

where

$$\xi_t^{HTM} = u'(c_t(\underline{a}, \underline{z})) - \partial_a \lambda_t(\underline{a}, \underline{z}) + \mu_t + \theta_t \frac{\epsilon_t}{\delta} \frac{d\tau_t(\underline{a}, \underline{z})}{dc_t(\underline{a}, \underline{z})} \quad \text{and} \quad \tilde{\chi}_t(a,z) = -u''(c_t(a,z)) \frac{\chi_t(a,z)}{g_t(a,z)}$$

as well as a set of initial conditions for the multipliers on forward-looking implementability conditions

$$0 = \theta_0 \tag{36}$$

$$0 = \phi_0(a,z). \tag{37}$$

The optimality conditions (30) – (32) hold everywhere in the interior of the idiosyncratic state space. For a formal treatment of boundary conditions, see Appendices B.2 through B.4

[26] In Appendix B.1, we first present a heuristic derivation of Proposition 4 in continuous time for the interior of the idiosyncratic state space. A formal treatment of boundary conditions follows in Appendices B.2 through B.4.

Equations (30) through (34) respectively correspond to the optimality conditions for i) the value function, ii) the cross-sectional distribution, iii) consumption, iv) aggregate labor, and v) wage inflation. Equation (35) corresponds to the optimality condition for the nominal interest rate.

The optimality conditions for the standard Ramsey problem (Proposition 4) can be seen as an augmented version of the optimality conditions for policy under discretion (Proposition 1). In particular, the multipliers $\chi_t(a,z)$, $\lambda_t(a,z)$, and μ_t, as well as ζ_t^{HTM}, have the same interpretation as in the discretion case—see pages 13 and 14. With commitment, the planner can counteract inflationary bias in the long run by making promises. These promises are encoded in the multipliers on the two forward-looking implementability conditions: θ_t for the Phillips curve and $\phi_t(a,z)$ for households' Bellman equations. Under discretion, these multipliers vanish. In fact, equations (31), (32), (33), and (35) correspond exactly to the optimality conditions (19) – (22) for policy under discretion when $\theta_t = 0$ and $\phi_t(a,z) = 0$.

The multiplier associated with the Phillips curve, θ_t, has the interpretation of a penalty (reward) associated with increasing inflation when $\theta_t > (<) \ 0$. Hence, we refer to θ_t as an *inflation penalty*.[27] Any perturbation that increases (decreases) the aggregate inflation-relevant labor wedge leads to deflationary (inflationary) pressure through the Phillips curve (see footnote 9). Changes in this labor wedge at date t are thus valued both directly and indirectly due to their effect on past inflation. The inflation penalty θ_t encodes the cumulative valuation of these resulting changes in past inflation. It thus appears in all perturbations that affect the aggregate inflation-relevant labor wedge: (i) The equation for social lifetime value (31) considers an increase in the mass of households in state (a,z). If the individual inflation-relevant labor wedge is positive (negative) for these states, $\tau_t(a,z) > (<) \ 0$, this perturbation puts negative (positive) pressure on inflation. (ii) In the social consumption-savings optimality condition (32), increasing consumption generates inflationary pressure since it reduces individual inflation-relevant labor wedges, $\frac{d\tau_t(a,z)}{dc_t(a,z)} < 0$. Finally (iii) in the aggregate activity condition (33), increasing hours worked also generates inflationary pressure by reducing individual inflation-relevant labor wedges, $\frac{d\tau_t(a,z)}{dN_t} < 0$.[28]

The multipliers associated with households' Bellman equation, $\phi_t(a,z)$, have the interpretation of a penalty (reward) associated with an increase in lifetime utility when $\phi_t(a,z) < (>) \ 0$. Hence, we

[27] As the multiplier on a forward-looking equation, θ_t encodes the impact on the Lagrangian (welfare) at time 0 from a change in inflation at time t. This is analogous to multipliers on backward-looking equations; for example, the multiplier on the capital accumulation equation in the neoclassical growth model encodes the present discounted value of a change in the capital stock. A change in inflation at time t affects inflation at all prior dates $s \in [0,t)$ through the forward-looking Phillips curve. These changes in inflation appear in the time-0 Lagrangian and are valued by the planner. If the planner were to reoptimize at time t, she would disregard these effects on past inflation. The inflation penalty θ_t encodes the associated welfare impact so that, if the planner reoptimizes at time t when confronted with θ_t, she will behave consistently with time-0 optimization. In summary, θ_t captures the cumulative, backward-looking impact on time-0 welfare resulting from a change in inflation at time t, appropriately discounted to period t.

[28] The inflation penalty θ_t also appears in the definition of ζ_t^{HTM}, the social valuation of increasing wealth for households in state $(\underline{a},\underline{z})$. As in the social consumption-savings condition, shifting wealth towards and thus increasing the consumption of hand-to-mouth households leads to inflationary pressure by lowering the aggregate inflation-relevant labor wedge.

refer to $\phi_t(a,z)$ as *distributional penalties*.[29] The distributional penalties appear in all perturbations that affect date-t lifetime values: (i) In the aggregate activity condition (33), increasing hours worked leads to a change in all households' flow utility that is captured by the individual labor wedges, $zu'(c_t(a,z)) - \frac{v'(N_t)}{A_t}$. (ii) And in the optimality condition for interest rates (35), the distributive pecuniary effect of a rate increase changes a household's utility by $a\partial_a V_t(a,z)$.

Relative to the optimality conditions for policy under discretion, Proposition 4 also features two new equations, (30) and (34). These define the laws of motion for $\phi_t(a,z)$ and θ_t. Equation (30) is central to this paper. It takes the form of a Kolmogorov forward equation augmented to account for births and deaths. This equation shows that the evolution of the distribution of distributional penalties $\phi_t(a,z)$ must be consistent with the evolution of households across idiosyncratic states, via $\mathcal{A}^*$. It also accounts for births of penalties, captured by the term $\partial_a \chi_t(a,z)$, as we explain in Section 4.4.

The optimality condition for inflation (34), which defines the law of motion for θ_t, simplifies to

$$\dot{\theta}_t = \delta \pi_t^w \tag{38}$$

whenever the distributional penalties add up to zero, $\iint \phi_t(a,z)\, da\, dz = 0$. We have proven that this is the case at a stationary Ramsey plan (defined below) and strongly conjecture that this condition holds at all times, which we have verified numerically.

4.2 Optimal Long-Run Policy

We start unpacking the implications of Proposition 4 by characterizing the optimal long-run policy under commitment. To that end, we define a stationary Ramsey plan, towards which a Ramsey plan may converge when all shocks converge as $t \to \infty$.

Definition 4. (Stationary Ramsey Plan) *A stationary Ramsey plan, with* $(A_t, \rho_t, \epsilon_t) = (A_{ss}, \rho_{ss}, \epsilon_{ss})$ *constant, is given by (i) an inflation rate,* π_{ss}^w*, aggregate hours,* N_{ss}*, stationary individual consumption allocations and value functions,* $c_{ss}(a,z)$ *and* $V_{ss}(a,z)$*, and a stationary cross-sectional distribution,* $g_{ss}(a,z)$*; (ii) a stationary Ramsey policy,* i_{ss}*; and (iii) a set of stationary multipliers,* $\phi_{ss}(a,z)$*,* $\lambda_{ss}(a,z)$*,* $\chi_{ss}(a,z)$*,* μ_{ss}*, and* θ_{ss}*, such that the optimality conditions and the implementability conditions for a Ramsey plan are satisfied.*

What are the implications of household heterogeneity for optimal long-run inflation? When policy

[29] Like θ_t, $\phi_t(a,z)$ is a multiplier on a forward-looking implementability condition. Any perturbation that shifts lifetime values $V_t(a,z)$ at date t affects past lifetime values $V_s(a,z)$, for $s \in [0,t)$, through the Bellman equation and, therefore, past consumption decisions $c_s(a,z)$ through the individual consumption-savings condition. The multiplier $\phi_t(a,z)$ encodes the cumulative, backward-looking social valuation of these indirect effects on past consumption that result from a change in value $V_t(a,z)$ at date t.

is set with discretion, the planner's redistribution motive substantially exacerbates inflationary bias. Proposition 5 shows that the stationary Ramsey plan features zero inflation even in the presence of household heterogeneity. Optimal policy under commitment therefore addresses the inflationary bias associated with discretion in the long run.[30]

Proposition 5. (**Optimal Long-Run Policy**) *With commitment, optimal long-run price inflation in both HANK and RANK is zero. That is,* $\pi_{\text{ss}} = \pi_{\text{ss}}^{\text{RA}} = 0$.

Our HANK model features long-run neutrality of monetary policy: In any competitive stationary equilibrium, the real interest rate and the allocation are pinned down by real forces. The only choice that the planner has is the split between nominal interest rate and nominal price inflation for a given real interest rate. Crucially, inflation and the nominal interest rate symmetrically affect households' financial income, which itself is proportional to the real interest rate $r_{\text{ss}} = i_{\text{ss}} - \pi_{\text{ss}}$. Therefore, since maintaining non-zero inflation is costly due to nominal rigidities while adjusting the nominal rate is not, the planner finds it optimal to exclusively use the nominal interest rate in the stationary Ramsey plan while promising to keep inflation at zero.[31] Formally, any stationary Ramsey plan must feature $\dot{\theta}_{\text{ss}} = 0$ since θ_{ss} is constant. Equation (38) then directly implies that optimal long-run inflation is zero, $\pi_{\text{ss}} = \pi_{\text{ss}}^{w} = 0$.

Importantly, the planner can only maintain zero inflation in the long run under commitment. As we discuss in Section 3, the planner always faces a time inconsistent incentive to overheat the economy, both to address the markup distortion and to redistribute toward indebted, high marginal utility households. Under discretion, this incentive is self-defeating as it simply results in inflationary bias. With commitment, the planner promises to keep inflation at zero in the long run in the absence of shocks.

4.3 Time Inconsistency, Timeless Penalty, and the Timeless Ramsey Problem

A planning problem is time inconsistent if the optimality conditions pinning down policy, allocation, and prices at some time t depend on the time at which the optimization takes place. The implementability conditions that constrain the Ramsey problem (28) include two sets of forward-looking conditions: individual Bellman equations and the New Keynesian Phillips curve. Each of these conditions is a source of time inconsistency.[32] While the standard Ramsey planner chooses

[30] Lemma 20 in the Appendix explicitly describes the conditions that characterize a stationary Ramsey plan.

[31] The fact that the optimal long-run policy features zero inflation should be understood as a benchmark. When inflation and the nominal interest rate have a differential impact across households, we may expect an optimal long-run policy that features non-zero inflation. Other frictions, such as those studied in Chari and Kehoe (1999), Khan et al. (2003), and Schmitt-Grohé and Uribe (2010), could also imply a non-zero optimal long-run inflation.

[32] The conditions under which forward-looking implementability conditions lead to time inconsistency in planning problems are well understood since Kydland and Prescott (1977).

policy with commitment from time 0 onwards, time inconsistency still manifests at time 0. This is often referred to as the "time-0 problem" (Kydland and Prescott, 1980).

Formally, the time-0 problem materializes as follows: The optimality conditions for the standard Ramsey problem require the initial conditions $\theta_0 = 0$ and $\phi_0(a,z) = 0$ for all (a,z) because initial inflation π_0^w and lifetime value $V_0(a,z)$ respectively are free. But any stationary Ramsey plan will generically feature $\theta_{ss} \neq 0$ and $\phi_{ss}(a,z) \neq 0$. This follows directly from the stationary version of equations (33) and (30). Intuitively, the standard Ramsey planner benefits from making promises for inflation and lifetime values in the long run; these are encoded in θ_{ss} and $\phi_{ss}(a,z)$. But at time 0, there are no such past promises. Hence, even if we initialize the economy at the allocation that obtains at the stationary Ramsey plan, i.e., $g_0(a,z) = g_{ss}(a,z)$, the planner will not set policy to $i_0 = i_{ss}$ in the absence of shocks, which would keep the economy at the stationary Ramsey plan. This would violate the initial conditions of the standard Ramsey problem, $\theta_0 = 0$ and $\phi_0(a,z) = 0$, and precisely formalizes the time-0 problem in our setting.

We illustrate the time-0 problem in Figure 1, which plots optimal inflation in the absence of shocks under different planning problems.[33] The solid and dashed red lines illustrate the inflationary bias associated with policy under discretion. The solid green line plots inflation under the standard Ramsey plan in the absence of shocks. Inflation converges to zero, thus addressing inflationary bias in the long run (Proposition 5). However, the time-0 problem implies that the standard Ramsey plan still features inflationary bias in the short run. As the world starts at time 0, no past promises constrain the planner, and so she finds it optimal to boost inefficiently low employment due to markups and redistribute towards indebted households, generating inflation in the short run.

Motivated by these observations, we now present a *timeless Ramsey approach* to resolve the time-0 problem. The associated timeless Ramsey plan—the solid blue line in Figure 1—resolves inflationary bias in both the short run and the long run. To that end, we introduce a particular time-0 penalty, the *timeless penalty*. Intuitively, confronting the Ramsey planner with the timeless penalty makes her internalize the promises she herself would like to make for the future. In Proposition 6, we show that the timeless penalty formalizes Woodford (1999)'s timeless perspective in our setting, so that the planner at time 0 behaves as if she had chosen policy with commitment infinitely long ago.

Definition 4. (**Time-0 Penalty**) *We define the time-0 penalty as*

$$\mathcal{T}(\phi,\theta) = \underbrace{\iint \phi(a,z) V_0(a,z)\, da\, dz}_{\text{Distributional Penalty}} \quad - \quad \underbrace{\theta \pi_0^w}_{\text{Inflation Penalty}} \tag{39}$$

[33] Figure 1 extends Figure 7.1 in Woodford (2003) and Figure 2 in Woodford (2010) to an environment with household heterogeneity.

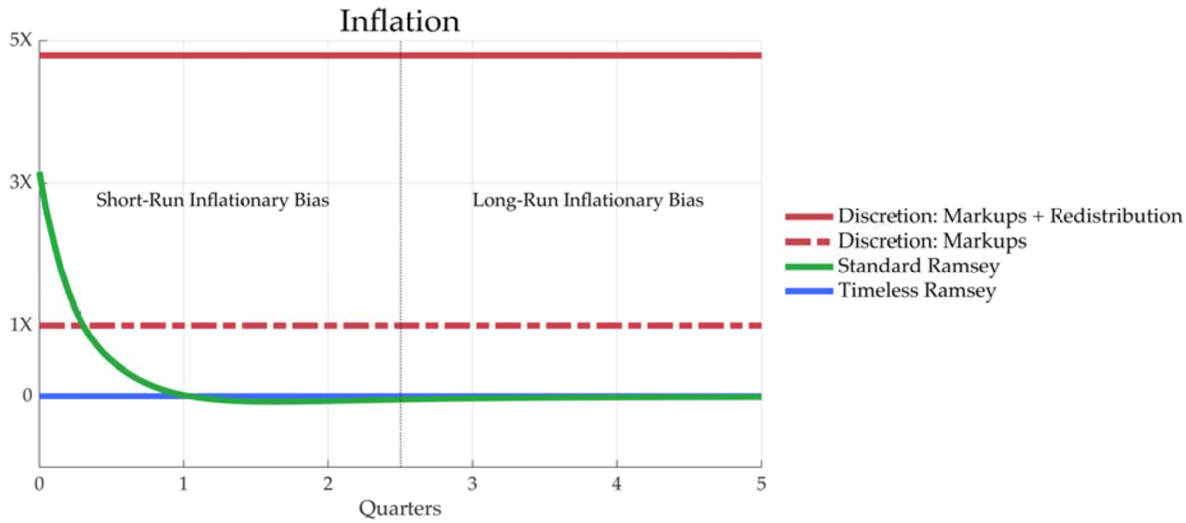


Figure 1. Inflationary Bias in the Short Run and the Long Run

Note. Figure 1 plots optimal inflation in the absence of shocks under different planning problems. We normalize to 1 the standard inflationary bias associated with monopolistic competition (red, dashed). The redistribution motive a utilitarian planner faces under discretion exacerbates long-run inflationary bias by a factor of 4 (red solid). Under the standard Ramsey problem (green), there is no inflationary bias in the long run. Due to the time-0 problem, however, there is still short-run inflationary bias. Only the timeless Ramsey problem (blue) resolves the time-0 problem and addresses inflationary bias in both the short run and the long run.

where we refer to $\phi(a,z)$ as a (per unit) distributional penalty and θ as a (per unit) inflation penalty.[34]

Building on Marcet and Marimon (2019), we introduce a penalty at time 0 for each forward-looking implementability condition that the Ramsey planner faces. We then define the *augmented Ramsey problem* in primal form as a modification of the standard Ramsey problem that confronts the planner with a time-0 penalty. Finally, we define the *timeless Ramsey problem* as the augmented Ramsey problem in which the time-0 penalty is chosen to resolve the time-0 problem. That is, a timeless Ramsey planner has no incentive to deviate from the stationary Ramsey plan at time 0 in the absence shocks.[35]

Definition 5.

(a) (**Augmented Ramsey Problem**) *The augmented Ramsey problem in primal form solves*

$$\min_{\{\phi_t(a,z),\, \chi_t(a,z), \lambda_t(a,z), \mu_t, \theta_t\}} \max_{\{c_t(a,z), V_t(a,z), g_t(a,z), N_t, \pi^w_t, i_t\}} L^{\mathrm{AP}}(g_0, \phi, \theta), \tag{40}$$

[34] To streamline the exposition, we use the term penalty to refer to i) $\mathcal{T}(\phi,\theta)$; ii) its components $\iint \phi(a,z)V_0(a,z)\,da\,dz$ and $\theta\pi^w_0$; and iii) the values of $\phi(a,z)$ and θ. The first three terms are total penalties, while the last two correspond to penalties per unit of lifetime utility and inflation, respectively.

[35] In Appendix B.6, we also characterize the dual form of the augmented and timeless Ramsey problems.

where $L^{\text{AP}}(g_0, \phi, \theta)$ denotes the augmented primal Lagrangian, given an initial distribution g_0 as well as initial penalties ϕ and θ. The augmented primal Lagrangian is defined as

$$L^{\text{AP}}(g_0, \phi, \theta) = L^{\text{SP}}(g_0) + \mathcal{T}(\phi, \theta) \tag{41}$$

where $L^{\text{SP}}(g_0)$ is the standard primal Lagrangian (29) and $\mathcal{T}(\phi, \theta)$ a time-0 penalty (39).

(b) (**Timeless Ramsey Problem**) *The timeless Ramsey problem in primal form is an augmented Ramsey problem in which the time-0 penalty takes the form $\mathcal{T}(\phi, \theta) = \mathcal{T}(\phi_{ss}, \theta_{ss})$. We refer to $\mathcal{T}(\phi_{ss}, \theta_{ss})$ as the timeless penalty, and we define the timeless primal Lagrangian as*

$$L^{\text{TP}} = L^{\text{AP}}(g_{ss}, \phi_{ss}, \theta_{ss}). \tag{42}$$

The Lagrangian of the augmented Ramsey problem $L^{\text{AP}}(g_0, \phi, \theta)$ is defined for arbitrary initial penalties ϕ and θ. For example, the augmented Ramsey problem nests the standard one when we set $\phi(a, z) = 0$ and $\theta = 0$, implying $L^{\text{AP}}(g_0, 0, 0) = L^{\text{SP}}(g_0)$. It will become clear in the following that only the timeless Ramsey problem, in which $\phi(a, z) = \phi_{\text{ss}}(a, z)$ and $\theta = \theta_{\text{ss}}$, resolves the time-0 problem. In that case, the timeless penalty encodes precisely the promises that the Ramsey planner would like to make in the long run, i.e., in the stationary Ramsey plan. Intuitively, the timeless penalty introduces an artificial cost that, on the margin, exactly offsets the marginal benefit of time-inconsistent deviations from the stationary Ramsey plan at time 0. Our approach shows that it is possible to transform the standard Ramsey problem into a timeless problem by simply augmenting it with the timeless penalty $\mathcal{T}(\phi_{ss}, \theta_{ss})$.[36]

Formally, a time-0 penalty enforces a new set of initial conditions on the two multipliers associated with forward-looking implementability conditions. In continuous time, the choice of initial lifetime values $V_0(a, z)$ and inflation π_0^w is free under the standard Ramsey problem, which gives rise to the initial conditions (36) and (37) of the standard Ramsey plan. Indeed, $\phi_0(a, z) = 0$ and $\theta_0 = 0$ is precisely an expression of the fact that the standard Ramsey planner is not bound by past promises on lifetime values and inflation at time 0, even though she would like to bind her future self by making such promises. By augmenting the Ramsey problem with the time-0 penalty

[36] Formally, a planner solving problem (42) sets policy at time 0 as if she had set policy with commitment infinitely far in the past. The timeless Ramsey plan associated with (42) corresponds exactly to optimal policy from a timeless perspective (Woodford, 1999). A timeless policy, as defined by Woodford (2010), represents a policy that

> "*even if not what the policy authority would choose if optimizing afresh at a given date t, [...] it should have been willing to commit itself to follow from that date t onward if the choice had been made at some indeterminate point in the past, when its choice would have internalized the consequences of the policy for expectations prior to date t.*"

$\mathcal{T}(\phi, \theta)$, we technically enforce new initial conditions,

$$\phi_0(a,z) = \phi(a,z) \tag{43}$$

$$\theta_0 = \theta, \tag{44}$$

which in turn constrain the planner's choice of initial lifetime values and inflation. The optimality conditions of the augmented Ramsey problem comprise exactly the same equations as in Proposition 4, i.e., equations (30) through (35), except that the initial conditions for the multipliers are now given by $\phi_0(a,z) = \phi(a,z)$ and $\theta_0 = \theta$. In economic terms, it is as if the penalties $\phi(a,z)$ and θ enforce artificial past promises. And when we initialize the time-0 penalty at $\phi = \phi_{\text{ss}}$ and $\theta = \theta_{\text{ss}}$, it is as if the timeless Ramsey planner is confronted with the same promises at time 0 that she herself would like to make in the long run.

Having introduced the timeless Ramsey problem, we now prove that it resolves the time-0 problem in Proposition 6, which is the main result of this subsection. Anticipating the sequence-space representation of our model (Section 5), we can interpret all endogenous variables as functions of the time paths of (i) policy, which we denote by $\boldsymbol{i} = \{i_t\}$, and (ii) exogenous shocks, which we denote by $\boldsymbol{Z} = \{A_t, \rho_t, \epsilon_t\}$, as well as (iii) initial conditions for the distribution $g_0(a,z)$ and penalties $(\phi(a,z), \theta)$. Given a sequence of shocks, an initial distribution, and initial promises, the planner then chooses among those competitive equilibria that are implementable by policy. In particular, we can evaluate the objective L^{TP} at any feasible policy path $\boldsymbol{i}$. A policy is then locally optimal when the derivative of L^{TP} with respect to any feasible perturbation of the policy path $d\boldsymbol{i}$ is 0.

Proposition 6. (**Timeless Ramsey Problem Resolves Time-**0 **Problem**) *Optimal policy under the timeless Lagrangian is time-consistent at the stationary Ramsey plan under the timeless penalty $\mathcal{T}(\phi_{ss}, \theta_{ss})$. That is,*

$$\frac{d}{d\boldsymbol{i}} L^{\text{TP}}\Big(g_{\text{ss}}, \phi_{\text{ss}}, \theta_{\text{ss}}, \boldsymbol{i}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}\Big) = 0. \tag{45}$$

Equation (45) says that, when we initialize the economy at the stationary Ramsey plan, i.e., $g_0(a,z) = g_{\text{ss}}(a,z)$, and set the time-0 penalty using the stationary multipliers, i.e., $\phi(a,z) = \phi_{\text{ss}}(a,z)$ and $\theta = \theta_{\text{ss}}$, then in the absence of shocks, i.e., $\boldsymbol{Z} = \boldsymbol{Z}_{\text{ss}}$, the stationary Ramsey policy is optimal, i.e., $\frac{d}{d\boldsymbol{i}} L^{\text{TP}}(\cdot) = 0$ when we set $\boldsymbol{i} = \boldsymbol{i}_{\text{ss}}$. Proposition 6 proves that a timeless Ramsey planner has no incentive to deviate from the stationary Ramsey plan in the absence of shocks. The timeless Ramsey problem resolves the time-0 problem and addresses inflationary bias in both the short run and the long run. The solid blue line in Figure 1 shows that the timeless Ramsey planner sets inflation to 0 in the absence of shocks.

Purpose of the timeless Ramsey approach. It should be evident that, from a time-0 perspective, the standard Ramsey plan attains higher welfare than the timeless Ramsey plan. The timeless Ramsey problem may thus be viewed as an inferior guide for policy design. However, there are at least three reasons why the timeless Ramsey approach is valuable. First, Woodford (1999)'s concerns about the time-0 problem remain valid: From a policymaker's perspective, access to new information and advances in modeling often necessitate a reevaluation of the framework used for policy design. If optimal policy is then recomputed each time under the standard Ramsey problem, it repeatedly suffers from the time-0 problem, which Woodford (1999) argues is an impractical guide for policy design. Second, the timeless Ramsey approach allows us to isolate the planner's pure stabilization motive in response to business cycle shocks and separate it from the time inconsistent incentive to deviate from the stationary Ramsey plan at time 0. We develop this argument in Sections 4.6 and 5. Third, we show in Section 5 that perturbation methods only yield valid approximations of optimal stabilization policy under the timeless Ramsey problem.

4.4 Properties of the Timeless Inflation and Distributional Penalties

We introduced the timeless penalty in Section 4.3 and showed that it resolves the time-0 problem, disincentivizing the planner from generating inflation in the short run. In this subsection, we explore this timeless penalty analytically. We establish two main results. First, we show that the inflation penalty, θ_{ss}, which is already present in RANK economies, depends on novel distributional considerations in HANK. Second, we show that the new distributional penalty that we introduce in this paper penalizes the welfare gains of indebted, high marginal utility households. The distributional penalty solves a novel promise-keeping Kolmogorov forward equation.

Timeless inflation penalty. Proposition 7 introduces an analytical characterization of the inflation penalty that resolves the time-0 problem in HANK. This expression combines the optimality conditions for consumption and hours worked with the Phillips curve.

Proposition 7. (**Timeless Inflation Penalty**) *The timeless penalty on inflation in both RANK and HANK economies satisfies*

$$\theta_{ss} = \frac{\Omega^1_{ss} + \Omega^2_{ss}}{-\frac{\epsilon}{\delta}\iint z\left(\frac{d\tau_{ss}(a,z)}{dc_{ss}(a,z)} + \frac{1}{A}\frac{d\tau_{ss}(a,z)}{dN_{ss}}\right) g_{ss}(a,z)\, da\, dz} \tag{46}$$

where Ω^1_{ss} and Ω^1_{ss} are given by

$$\Omega^1_{ss} = \overbrace{\left(1 - \frac{\epsilon - 1}{\epsilon}(1 + \tau^L)\right)}^{=0 \textit{ with appropriate employment subsidy}} \iint z u'(c_{ss}(a,z)) g_{ss}(a,z)\, da\, dz$$

$$\Omega^2_{ss} = \underbrace{\iint \left(z u'(c_{ss}(a,z)) - \frac{v'(n_{ss})}{A}\right)\phi_{ss}(a,z)\, da\, dz - \iint z \tilde{\chi}_{ss}(a,z) g_{ss}(a,z)\, da\, dz}_{=0 \textit{ in RANK}}$$

Ω^1_{ss} is 0 under the appropriate employment subsidy and Ω^2_{ss} is 0 in the RANK limit. Since Ω^2_{ss} is typically non-zero, distributional considerations shape the inflation penalty in HANK economies.

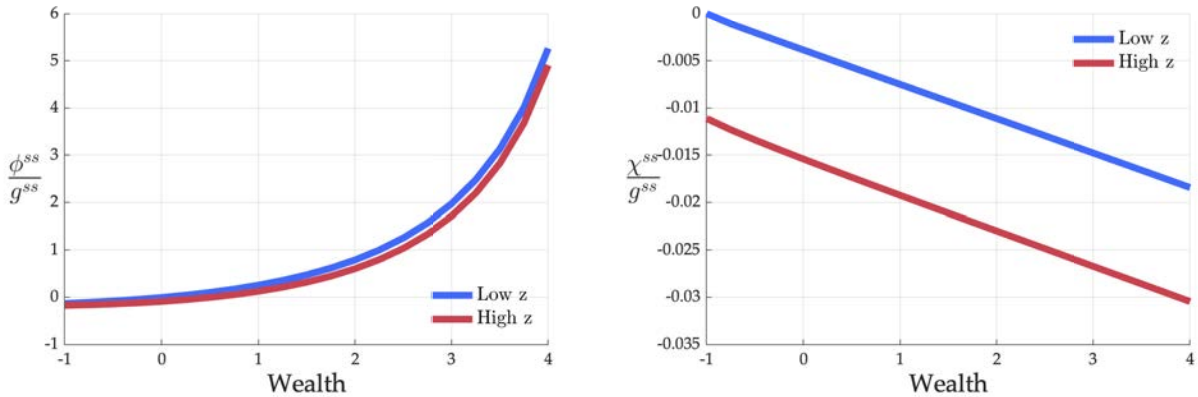
As we show in the Appendix, the denominator of equation (46) is always positive, so the sign of the inflation penalty depends on the signs of Ω^1_{ss} and Ω^2_{ss}. In the RANK limit, distributional considerations disappear and Ω^2_{ss} vanishes. In this case, the inflation penalty inherits the sign of $(1 - \frac{\epsilon-1}{\epsilon}(1 + \tau^L))$. In RANK, time inconsistency only emerges when employment is inefficiently low due to markups in a distorted steady state. In fact, with the appropriate employment subsidy, $\frac{\epsilon-1}{\epsilon}(1 + \tau^L) = 1$, equation (46) implies that no inflation penalty is required because no time consistency problem emerges in that case.

These same forces that shape the inflation penalty in RANK also appear in our HANK economy. However, the inflation penalty in HANK is also shaped by distributional considerations. Even with the correct employment subsidy to address the markup distortion in steady state, the time consistency problem on inflation does not disappear. Intuitively, θ_{ss} impacts the planner's desire to perturb aggregate economic activity by penalizing inflation at time 0. When households are heterogeneous, changes in aggregate economic activity have distributional consequences. In particular, the first term in Ω^2_{ss} captures the differential impact of an increase in hours worked on households' flow utility, while the second term accounts for the differential impact on households' consumption-savings decisions.

In other words, the two sources of time inconsistency—markups and redistribution—meaningfully interact now. A corollary of this result is that the choice of an appropriate inflation penalty takes on a distributional dimension whenever the planner has a utilitarian objective.

Timeless distributional penalty. In a RANK economy, the nominal interest rate is sufficient to fully correct the representative household's consumption-savings decision. Formally, we show that the Ramsey planner in RANK sets $\phi^{\mathrm{RA}}_t = 0$ at all times—see equation (84) in the Appendix. This implies that no time consistency problem separately emerges from the Bellman equation.

With heterogeneous households, the policy instrument still corrects households' decisions, but



(a) Distributional penalties

(b) Consumption-savings multiplier

Figure 2. Timeless Distributional Penalty

Note. The left panel of Figure 2 shows the steady state values of the timeless distributional penalties, $\phi_{ss}(a,z)$, normalized by the mass of households, $g_{ss}(a,z)$. The right panel of Figure 2 shows the steady state values of the multiplier on households' consumption-savings optimality condition, $\chi_{ss}(a,z)$, also normalized by the mass of households, $g_{ss}(a,z)$. See Section 5 for the calibration.

only on average, implying $\int\int \phi_t(a,z)\, da\, dz = 0$. Unlike in RANK economies, the nominal interest rate is no longer sufficient to correct the entire cross section of consumption-savings decisions. The Ramsey planner consequently finds that households privately consume too much or too little. Under commitment, the planner then finds it valuable to make promises about the future in order to influence consumption allocations today. These promises—encoded in the time-varying multiplier $\phi_t(a,z)$—open the door to time inconsistency. Our next result characterizes the timeless distributional penalty $\phi_{ss}(a,z)$ that addresses this time consistency problem.[37]

Proposition 8. (**Timeless Distributional Penalty**) *The timeless distributional penalty $\phi_{ss}(a,z)$ solves the promise-keeping Kolmogorov forward equation*

$$0 = -\mathcal{A}^*_{ss}\phi_{ss}(a,z) + \partial_a \chi_{ss}(a,z), \tag{47}$$

where $\mathcal{A}^*_{ss}$ is the Kolmogorov forward operator associated with the stationary cross-sectional distribution.

Proposition 8 shows that the multipliers on households' Bellman equations are themselves characterized by Kolmogorov forward equations. We refer to equation (30) and its stationary counterpart (47) as *promise-keeping Kolmogorov forward* equations because they characterizes the evolution of

[37] Chari et al. (2020) study optimal capital taxation with heterogeneous agents. They address the time-0 problem by augmented the Ramsey problem with a "wealth constraint" that prevents taxing away free rents. Our penalty performs a similar function but its form derives from the recursive multiplier approach of Marcet and Marimon (2019).

the planner's promises on individual lifetime values. These promises, which also have the interpretation of penalties, as explained above, are encoded in the multiplier $\phi_{\text{ss}}(a,z)$. The evolution of distributional penalties in the cross section must be consistent with the law of motion of households, which is why the Kolmogorov forward operator $\mathcal{A}^*_{\text{ss}}$ appears in equation (47). Intuitively, penalties are associated with households in a given state, and so if a household transitions from one individual state to another, the penalty moves with her.[38] If individuals' optimal consumption-savings decisions change as they transition between different wealth states, then $\partial_a \chi_t(a,z)$ can be interpreted as capturing "births" and "deaths" of relative promises in the cross section. The reason why $\partial_a \chi_t(a,z)$ enters the promise-keeping Kolmogorov equations is because changes in lifetime utility for a given state (a,z) impact the consumption of households at a and slightly above (or below), per the households' consumption-savings optimality condition. Solving for $\phi_t(a,z)$ and $\chi_t(a,z)$ jointly and characterizing how they are linked via this promise-keeping Kolmogorov forward equation is one of the contributions of this paper.[39]

Figure 2 illustrates the distributional penalty $\phi_{\text{ss}}(a,z)$. Panel (a) shows that $\phi_{\text{ss}}(a,z) < 0$ for indebted households, which is consistent with an interpretation of $\phi_{\text{ss}}(a,z)$ as a penalty on redistribution. Intuitively, in order to counteract the time-inconsistent incentive to redistribute towards high marginal utility households, the planner with commitment sets distributional penalties that penalize welfare assessments that benefit such households. While it may seem counterintuitive that a utilitarian planner penalizes high marginal utility, this is the way in which a planner with commitment fights the inflationary bias present under discretion. Panel (b) of Figure 2 displays the stationary consumption-savings multiplier, $\chi_{\text{ss}}(a,z)$, normalized by the mass of households, $g_{\text{ss}}(a,z)$. In this particular calibration, $\chi_{\text{ss}}(a,z) < 0$ for all households, which implies that the planner, if unconstrained, would like households to consume less. This is consistent with the fact that $\mu_{ss} < 0$ at the stationary Ramsey plan of our calibrated model.

4.5 Penalties vs. Targets

Sections 4.3 and 4.4 focus on resolving the time-0 problem that emerges in the standard Ramsey problem. To that end, we introduce the timeless penalty and study the timeless Ramsey problem (42). Implementing the resulting Ramsey plan still relies on an infinite sequence of promises, however, which may be unrealistic in practice. We now explore whether a central bank that sets policy under discretion can still implement the optimal commitment solution under an appropriate institutional arrangement or with the appropriate penalties or targets (Clarida et al., 1999).

[38] Note that the operator $\mathcal{A}^*_t$ is mass-preserving, i.e., $\iint \mathcal{A}^*_t \phi_t(a,z)\, da\, dz = 0$, which allows us to interpret $\phi_t(a,z)$ as a distribution (of penalties).

[39] We conjecture that promise-keeping Kolmogorov forward equations will appear in other models in which a continuum of Bellman equations act as constraints. We also conjecture that if households had additional margins of adjustment besides consumption-savings, new birth and death terms would augment the Kolmogorov equation.

Time-varying penalties. First, consider again problem (17), where a Ramsey planner sets policy with commitment over a finite horizon. In Section 3, we identify policy under discretion with the limit of this problem as the commitment horizon becomes vanishingly small, i.e., as $\psi \to \infty$. Leveraging the observation that a timeless penalty can resolve the time-0 problem of a standard Ramsey planner, we now confront each finite-horizon Ramsey planner with a penalty at the time of transition, just like we confronted the standard Ramsey planner with a timeless penalty at time 0. Intuitively, this sequence of timeless penalties ensures that each successive finite-horizon planner behaves as if she had committed to policy in the infinite past. The resulting sequence problem is given by

$$\tilde{\mathcal{W}}_0(g_0, \phi_0, \theta_0) = \min_{\boldsymbol{M}} \max_{\boldsymbol{X}} \mathbb{E}_0 \Big[\underbrace{L(0, \tau_1, g_0) + \mathcal{T}_0(\phi_0, \theta_0)}_{\text{Timeless Lagrangian: } L^{\text{TP}}} + e^{-\int_0^{\tau_1} \rho_s ds} \tilde{\mathcal{W}}_{\tau_1}(g_{\tau_1}, \phi_{\tau_1}, \theta_{\tau_1}) \Big], \tag{48}$$

where $\boldsymbol{M}$, $\boldsymbol{X}$, and $L(0, T, g_0)$ are defined as in (17). Crucially, the evolution of ϕ_t and θ_t is given by equations (30) and (34), and we initialize the timeless penalties at $\phi_0(a,z) = \phi_{\text{ss}}(a,z)$ and $\theta_0 = \theta_{\text{ss}}$. By modifying the flow payoff in equation (48), we confront each planner with the appropriate timeless penalty at the time of transition. And in the limit as $\psi \to \infty$, where planners transition instantaneously, the penalties are also "active" in every instant. In the discrete-time analysis of, e.g., Galí (2015), we would say that the Markov planner faces these penalties in every period.

The timeless Ramsey plan can be implemented under discretion as long as the planner (central bank) faces the appropriate time-varying penalty $\mathcal{T}_t(\phi_t, \theta_t)$, which includes an inflation penalty $-\theta_t \pi_t^w$ and a distributional penalty $\iint \phi_t(a,z) V_t(a,z)\, da\, dz$. Intuitively, the difference between the commitment and the discretion solutions are the multipliers associated with the forward-looking constraints. By modifying the Markov planner's flow utility to account for these terms in the form of time-varying penalties, it is possible to implement the commitment solution under discretion (Svensson, 1997; Marcet and Marimon, 2019; Clayton and Schaab, 2022).

Inflation targeting. While confronting the discretionary planner (central bank) with the time-varying penalty $\mathcal{T}_t(\phi_t, \theta_t)$ can implement the timeless Ramsey solution, central bank design in practice is commonly based on targeting frameworks.

Proposition 5 highlights that a strict zero-inflation target implements the Ramsey plan in the absence of shocks.[40] In other words, household heterogeneity in our environment does not alter the longstanding view that an inflation target can successfully resolve inflationary bias in steady state. What is surprising is that an implementation of such an inflation target based on penalties requires

[40] Formally, a Markov planner will implement the optimal stationary Ramsey plan in the absence of shocks when confronted either with an infinite penalty for non-zero inflation or with an additional implementability condition (constraint) that requires $\pi_t = 0$.

two distinct penalties in our setting, whereas only the inflation penalty is required in RANK economies. Both penalties and targets can be used to implement a particular solution because of their duality relation: Constraints (targets) on optimal policy problems can be transformed into costs (penalties) in the objective function, and vice versa.[41] However, since the planner has a single aggregate instrument, once a path of aggregate variables (in this case inflation) is used as a target, the choice of instrument is automatically determined.

In the presence of shocks, our paper demonstrates that flexible inflation targeting is in principle still the appropriate framework for policy design. This target would be anchored around zero inflation in our setting. And in response to a shock, the target would prescribe the path of inflation that is optimal under the timeless Ramsey plan. Hence, an important takeaway of our analysis is that optimal policy in our HANK economy can also be implemented by a flexible inflation target around zero inflation, where the flexibility to stabilize business cycle shocks is now governed by distributional considerations as in our Ramsey problem.

4.6 Optimal Stabilization Policy

Characterizing optimal stabilization policy under the standard Ramsey plan would conflate the pure stabilization motive of policy with the time-0 problem, i.e., the planner's time inconsistent incentive to deviate from the stationary Ramsey plan even in the absence of shocks. A key motivation for setting up the timeless Ramsey problem is that it isolates the pure stabilization motive by resolving the time-0 problem. In this final subsection, we study optimal stabilization policy under commitment, we focus on the timeless Ramsey plan.

In Proposition 9 we characterize a non-linear, exact targeting rule for optimal monetary policy with commitment in response to demand, productivity, and cost-push shocks.[42] By considering special cases of this targeting rule it is possible to recover i) the discretion targeting rule introduced in Proposition 2, which allows us to highlight the role of inflation and distributional penalties counteracting the forces that drive discretionary policy, and ii) the well understood targeting rule in RANK, which allows us to identify the implications of household heterogeneity for optimal stabilization.

Proposition 9. (**Targeting Rule for Stabilization Policy under Commitment**) *Optimal monetary*

[41] A planner could conceivably target instead a particular path of lifetime utilities. This approach, which we do not explore in our paper, connects our results to the work on recursive contracting, as in Sannikov (2008) and Williams (2011), among many others.

[42] This targeting, as well as the targeting rule under discretion in Proposition 2, can be interpreted as a double perturbation in which the planner makes all households work an additional hour and consume the output generated, while also increasing interest rates to neutralize the intertemporal impact of the perturbation.

stabilization policy is summarized by the targeting rule

$$0 = \underbrace{\iint \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) g_t(a,z)\, da\, dz}_{\text{Aggregate Labor Wedge}} - \underbrace{\Omega_t^D \iint a u'(c_t(a,z)) g_t(a,z)\, da\, dz}_{\text{Redistribution Motive}} \tag{49}$$

$$+ \underbrace{\iint (z - \Omega_t^D a) u'(c_t(a,z)) \phi_t(a,z)\, da\, dz}_{\text{Distributional Penalty}} + \underbrace{\theta_t \frac{\epsilon_t}{\delta} \iint \left(\frac{z}{A_t} \frac{d\tau_t(a,z)}{dN_t} + (z - \Omega_t^D a) \mathcal{M}_t(a,z) \frac{d\tau_t(a,z)}{dc_t(a,z)} \right) g_t(a,z)\, da\, dz}_{\text{Inflation Penalty}}$$

where Ω_t^D is defined in Proposition 2.

The first and second terms in equation (49) respectively correspond to the aggregate labor wedge and the distributive pecuniary effect of interest rate changes. The third and fourth terms correspond to the distributional and inflation penalties. We leverage this equation to present four results.

First, note that when $\phi_t(a,z) = 0$ and $\theta_t = 0$, equation (49) collapses to the targeting rule under discretion introduced in Proposition 2. Intuitively, the new terms that shape the targeting rule under commitment act as penalties for the planner, counteracting the forces that drive discretionary policy.

Second, the targeting rule (49) also allows us to revisit optimal monetary stabilization policy in RANK, which it nests. In RANK, $\phi_t(a,z) = a = 0$, so the targeting rule simply trades off aggregate stabilization—encoded in the aggregate labor wedge—with an inflation penalty. Suppose we allow for the appropriate steady state employment subsidy, so that $\frac{\epsilon-1}{\epsilon}(1+\tau^L) = 1$. If we consider demand and TFP shocks, $\pi_t = \theta_t = 0$ are always feasible, so the targeting rule (49) implies that the aggregate labor wedge is zero: In response to demand and TFP shocks, the Ramsey planner in RANK closes both the inflation and output gaps at all times. This is the Divince Coincidence benchmark (Blanchard and Galí, 2007). In RANK, there is no tradeoff between inflation and output in the absence of cost-push shocks. In the case of cost-push shocks, Divine Coincidence breaks down, even in RANK. From the Phillips curve, it follows that at zero inflation, the only way in which the aggregate inflation-relevant labor wedge is zero, is when the aggregate labor wedge is non-zero. The planner consequently cannot close the inflation and output gaps at the same time.

Third, in a HANK economy, the targeting rule for aggregate stabilization policy is shaped by distributional considerations. Hence, even though it is feasible for the planner to close the inflation and output gaps at the same time in the absence of cost-push shocks, she finds it optimal not to do so. Divine Coincidence consequently fails even with the appropriate employment subsidy and in the absence of cost-push shocks. Formally, the aggregate labor wedge that makes equation (49) hold in response to a shock need not be zero. While pinpointing the source of departure from Divine Coincidence in general is difficult, we present a quantitative decomposition in Section 5.2.

We show in Appendix B.10 that the targeting rule (49) admits an alternative representation that augments the discretionary output gap targeting rule (26) with two penalty wedges. Using this decomposition, we demonstrate in Section 5.2 that departures from Divine Coincidence in response to demand shocks are due to changes in the redistribution wedge.

Finally, optimal stabilization policy under the timeless Ramsey plan always features inflation overshooting. This result applies to both RANK and HANK economies and follows from the fact that the inflation penalty has initial and terminal conditions $\theta_0 = \lim_{T\to\infty} \theta_T = \theta_{ss}$. Its evolution in response to a shock is characterized by $\dot{\theta}_t = \delta \pi_t^w$. Hence, integrating and using the boundary conditions, it must be the case that

$$\int_0^\infty \pi_t^w dt = 0 \tag{50}$$

in response to any shock. That is, if inflation is positive on impact in response to a shock, $\pi_0^w > 0$, it must turn negative at some point in the future, and vice versa if $\pi_0^w < 0$.

5 Quantitative Analysis in Sequence Space

In this section, we extend the sequence-space approach (Boppart et al., 2018; Auclert et al., 2021) to Ramsey problems and welfare analysis. This allows us to compute transition dynamics under optimal policy—under discretion and with commitment—both non-linearly and using perturbation methods. We extend the fake-news algorithm of (Auclert et al., 2021) to compute optimal policy and show that the timeless Ramsey approach of Section 4 is critical for the validity of sequence-space perturbations.[43]

5.1 Sequence-Space Methods for Optimal Policy in HANK

We work with an abstract sequence-space representation of our model. Competitive equilibrium can be summarized by an *equilibrium map* that take as inputs the time paths of aggregates,

$$\mathcal{H}(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z}) = 0, \tag{51}$$

where $\boldsymbol{i} = \{i_t\}_{t\geq 0}$ denotes the path of policy, $\boldsymbol{Z} = \{A_t, \rho_t, \epsilon_t\}_{t\geq 0}$ the path of exogenous shocks, and $\boldsymbol{X}$ the path of macroeconomic aggregates. Given an initial cross-sectional distribution $g_0(a,z)$, which is implicitly encoded in $\mathcal{H}(\cdot)$, the equilibrium map (51) characterizes macroeconomic aggregates in terms of policy $\boldsymbol{i}$ and shocks $\boldsymbol{Z}$, i.e., $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z})$. The sequence-space representation (51) is as in

[43] Our perturbation approach is closest to that of Khan et al. (2003) and Schmitt-Grohé and Uribe (2004a) who also first characterize the optimality conditions that define a Ramsey plan non-linearly and then approximate these. It is well understood that, at least in the standard model, alternative valid perturbation methods also include the linear-quadratic approach (Benigno and Woodford, 2012) and evaluating welfare under a higher-order approximation of the equilibrium conditions (Schmitt-Grohé and Uribe, 2004b, 2007).

Auclert et al. (2021), except that $\mathcal{H}(\cdot)$ here also takes the path of policy $\boldsymbol{i}$ as an input, which is set optimally by the planner. Optimal policy, in turn, is determined as part of a Ramsey plan, whose sequence-space representation we characterize next.

Proposition 10. (**Sequence-Space Representation of Ramsey Plans**) *Given an initial distribution $g_0(a,z)$, initial penalties $\phi(a,z)$ and θ, as well as a path for exogenous shocks $\boldsymbol{Z}$, a timeless Ramsey plan comprises aggregate allocations and prices $\boldsymbol{X}$, optimal policy $\boldsymbol{i}$, and multipliers $\boldsymbol{M}$. Its sequence-space representation is given by*

$$\mathcal{R}(\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i}, \boldsymbol{Z}) = 0, \tag{52}$$

where we leave implicit the dependence of the Ramsey map $\mathcal{R}(\cdot)$ on $g_0(a,z)$, $\phi(a,z)$, and θ.

We prove the sequence-space representations of equilibrium (51) and Ramsey plans (52) in Appendices D.1 and D.2.

Our sequence-space representation of Ramsey plans is valid for any initial distribution $g_0(a,z)$ and initial penalties $\phi(a,z)$ and θ. Equation (52) therefore recovers the standard Ramsey plan of Proposition 4 when we set $\phi(a,z) = 0$ for all (a,z) and $\theta = 0$. Similarly, it follows from Proposition 6 that evaluating the Ramsey plan (52) at $(g_{\text{ss}}, \phi_{\text{ss}}, \theta_{\text{ss}})$ resolves the time-0 problem. In that case, we refer to it as a timeless Ramsey plan. In the following, we always initialize the penalties at $\phi(a,z) = \phi_{\text{ss}}(a,z)$ and $\theta = \theta_{\text{ss}}$, and focus on characterizing the response of optimal policy, $d\boldsymbol{i}$, to exogenous shocks, $d\boldsymbol{Z}$, under the timeless Ramsey plan.

The Ramsey plan representation (52) consists of two sets of equations. The first block is the system of equations (51), which characterizes aggregate allocations and prices $\boldsymbol{X}$ given policy $\boldsymbol{i}$ and shocks $\boldsymbol{Z}$. The second block comprises the first-order optimality conditions of the Ramsey problem that solve for aggregate multipliers $\boldsymbol{M}$ and policy $\boldsymbol{i}$. Crucially, the Ramsey equations that characterize optimal policy are coupled with those that describe the evolution of multipliers. Unlike the equilibrium map $\mathcal{H}(\cdot)$, which suffices to solve for transition dynamics given policy, the Ramsey map $\mathcal{R}(\cdot)$ takes as inputs the aggregate multipliers $\boldsymbol{M}$ and features the equations that characterize them.

In this sequence-space representation, we refer to a (timeless) Ramsey plan as the time paths of aggregates, $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i})$.[44] The system of equations (52) characterizes a Ramsey plan as a

[44] In Section 4.1, we defined a Ramsey plan as the time paths of both aggregates and individual objects—namely, individual allocations, the cross-sectional distribution, and individual multipliers. In the sequence-space representation of our economy, we can express these individual objects as functions of the time paths of aggregates, as we formally show in Appendices D.1 and D.2. We thus loosely refer to a Ramsey plan in sequence-space form as the time paths of aggregates $(\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i})$ with the understanding that the remaining individual objects can be expressed and easily obtained as functions of these.

function of the exogenous shocks, i.e.,

$$\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{Z}),$$

implicitly taking as given an initial distribution $g_0(a,z)$ as well as initial penalties $\phi(a,z)$ and θ. The sequence-space representation of Ramsey plans in Proposition 10 is not unique. One minimal representation of our baseline economy, which we use in our numerical implementation, is $\boldsymbol{X} = \{\Lambda_t, N_t\}_{t\geq 0}$, $\boldsymbol{M} = \{\mu_t, \theta_t\}_{t\geq 0}$, and $\boldsymbol{i} = \{i_t\}_{t\geq 0}$, where Λ_t is the aggregate labor wedge. In that case, the Ramsey plan representation (52) becomes a system of five equations: the definition of Λ_t as the aggregate labor wedge, the resource constraint (16), as well as the three aggregate optimality conditions (33), (34), and (35). Together, they solve for the Ramsey plan as a function of shocks, i.e., $\boldsymbol{X}(\boldsymbol{Z})$, $\boldsymbol{M}(\boldsymbol{Z})$, and $\boldsymbol{i}(\boldsymbol{Z})$, taking as given an initial cross-sectional distribution $g_0(a,z)$, as well as initial penalties $\phi(a,z)$ and θ.

5.1.1 Non-Linear Optimal Policy

The sequence-space representation of Ramsey plans in Proposition 10 is a system of non-linear equations. We can directly solve (52) non-linearly to compute optimal stabilization policy around the stationary Ramsey plan for any sequence of shocks $\boldsymbol{Z}$ that reverts back to $\boldsymbol{Z}_{\text{ss}}$. Computing the timeless Ramsey plan non-linearly is tractable and fast in our baseline HANK economy. Using an efficient quasi-Newton algorithm, we can solve (52) non-linearly in less than 10 seconds.[45]

However, computing non-linear transition paths in more complex HANK economies with richer cross-sectional heterogeneity can become cumbersome. Local perturbation methods, on the other hand, are fast and oftentimes very accurate in the context of canonical HANK environments.[46] In the remainder of this section, we develop sequence-space perturbation methods to approximate optimal policy in a neighborhood around the stationary Ramsey plan. In principle, we can take either the primal or the dual representation of our Ramsey problem as a starting point to approximate optimal policy. In Sections 5.1.2 and 5.1.3, we present both approaches and argue that they have distinct advantages and disadvantages in different contexts.

[45] We use the quasi-Newton algorithm developed by Schaab and Zhang (2022) and Schaab (2020) to compute non-linear transition paths in heterogenous-agent economies. The code is available at `https://github.com/schaab-lab/SparseEcon`. Using this solver, computing the non-linear Ramsey plan of our model takes less than 10 seconds on a personal computer for discretized time grids with 150 nodes, using a 2020 13-inch MacBook Pro with an M1 chip and 16 GB memory.

[46] When computing Ramsey plans non-linearly, we use quasi-Newton rather than standard Newton methods. This means that we compute the Jacobians involved in the algorithm once and subsequently use a recursive approximation. In practice, the algorithm never has to recompute the Jacobians and converges quickly, precisely because first-order perturbation solutions are typically very accurate approximations in canonical HANK economies. Therefore, the objects we need are precisely those we also compute below in Section 5.1.2, i.e., $\mathcal{R}_{\boldsymbol{R}}$ and $\mathcal{R}_{\boldsymbol{Z}}$ evaluated around the stationary Ramsey plan, using a fake-news algorithm. As long as the quasi-Newton algorithm does not require that we recompute the Jacobian matrix, computing the non-linear solution is just as fast as the fake-news algorithm for the perturbation approach, requiring the computation of only a single column of the Jacobians $\mathcal{R}_{\boldsymbol{R}}$ and $\mathcal{R}_{\boldsymbol{Z}}$.

5.1.2 Optimal Policy Perturbations in the Primal

To approximate optimal policy in the primal representation of the Ramsey problem, we take as our starting point the system of equations (52).

Proposition 11. (**Optimal Policy Perturbations in the Primal**) *Consider the primal Ramsey problem and the associated Ramsey plan, which is characterized by* (52) *and solves* $\mathcal{R}(\cdot) = 0$*. Suppose we initialize the Ramsey plan at the cross-sectional distribution* $g_0(a, z) = g_{ss}(a, z)$ *and with initial timeless penalties* $\phi(a, z) = \phi_{ss}(a, z)$ *and* $\theta = \theta_{ss}$*. To first order, optimal stabilization policy is then characterized as part of the timeless Ramsey plan by*

$$d\boldsymbol{R} = -\mathcal{R}_{\boldsymbol{R}}^{-1}\mathcal{R}_{\boldsymbol{Z}} d\boldsymbol{Z} \tag{53}$$

where $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{ss}$ *is the exogenous shock,* $d\boldsymbol{R} = (d\boldsymbol{X}, d\boldsymbol{M}, d\boldsymbol{i})$ *denotes the response of the Ramsey plan, and* $\mathcal{R}_{\boldsymbol{R}}$ *and* $\mathcal{R}_{\boldsymbol{Z}}$ *are Jacobians of the Ramsey plan map.*

We prove Proposition 11 in Appendix D.3.

It is critical to note that the validity of the sequence-space perturbation method in Proposition 11 relies on initializing the Ramsey problem with the timeless penalties, so that $\mathcal{R}(\cdot)$ characterizes a timeless Ramsey plan. With the timeless penalties, $d\boldsymbol{R}$ only captures the planner's stabilization motive in response to shocks $d\boldsymbol{Z}$. Without them, $d\boldsymbol{R}$ conflates the stabilization motive with the time-0 problem and is consequently no valid solution of optimal stabilization policy to first order. Our timeless Ramsey approach is therefore the critical foundation that allows us to leverage perturbation methods to compute optimal stabilization policy.

To approximate Ramsey plans to first order in the primal, we have to compute two first-order derivative matrices, $\mathcal{R}_{\boldsymbol{R}}$ and $\mathcal{R}_{\boldsymbol{Z}}$. These matrices can in turn be constructed from sequence-space Jacobians, which allows us to leverage the power of sequence-space perturbation methods. In Appendix D, we extend the fake-news algorithm developed by Auclert et al. (2021) for sequence-space Jacobians to compute optimal policy via the Ramsey map Jacobians $\mathcal{R}_{\boldsymbol{R}}$ and $\mathcal{R}_{\boldsymbol{Z}}$.[47]

5.1.3 Optimal Policy Perturbations in the Dual

Appendix B.6 formally introduces the dual form of our timeless Ramsey problem. While the previous subsection directly uses the primal representation of Ramsey plans, an alternative sequence-

[47] Auclert et al. (2021) show how to use the equilibrium map (51) to efficiently compute transition dynamics for a given path of policy to first order. They develop a general model representation of the standard micro block of competitive equilibria in heterogeneous-agent economies, i.e., the set of equations that characterize the allocations and behavior of individual agents. We show in Appendix D that computing optimal policy using the sequence-space Ramsey plan representation (52) requires a second "micro block," namely the set of individual multiplier equations. We develop a general sequence-space representation for this multiplier block and show that the same principles underlying Auclert et al. (2021)'s fake-news algorithm can be used to efficiently compute sequence-space Jacobians for multipliers.

space perturbation method can be developed by using the dual form as a starting point. We do so in Appendix D.4.

The key advantage of the dual approach is that multipliers do not explicitly have to be computed as part of the Ramsey plan solution. When the multiplier equations are particularly complex and computationally intensive, this can be an important advantage. The main disadvantage of the dual approach is that it relies on second-order derivatives, whereas the primal approach relies on first-order derivatives. In Section D, we therefore introduce *sequence-space Hessians* as the natural second-order generalization of sequence-space Jacobians. Finally, Appendix D.4 offers a detailed discussion on the advantages and disadvantages of the dual approach relative to the primal approach of Section 5.1.2.

5.2 Optimal Stabilization Policy: Quantitative Analysis

We now compute optimal monetary stabilization policy in response to demand shocks (this section), as well as TFP (Appendix F.1) and cost-push (Appendix F.2) shocks.

Calibration. Adopting isoelastic preferences, we set the discount rate to a quarterly $\rho = 0.02$, the elasticity of intertemporal substitution to $\gamma = 2$, and the inverse Frisch elasticity to $\eta = 2$. We set the elasticity of substitution between labor varieties to $\epsilon = 10$ and the nominal wage adjustment cost to $\delta = 100$, following standard practice in the wage rigidity literature (Auclert et al., 2020). Finally, we allow for an employment subsidy $(1 + \tau^L)\frac{\epsilon-1}{\epsilon} = 1$ that offsets the wage-markup distortion in stationary equilibrium.

In our HANK model, we model households' earnings risk as a two-state Markov chain with $z_t \in \{\underline{z}, \overline{z}\}$, where $\underline{z} = 0.8$ and $\overline{z} = 1.2$. We set the quarterly Poisson transition rate out of both states to 0.33. Our RANK benchmark can be seen as the limit as $\underline{z}, \overline{z} \to 1$, using as initial condition for the cross-sectional distribution a Dirac mass at $(a, z) = (0, 1)$.

Finally, we model the demand shock as a mean-reverting AR(1) process. In continuous time, this implies that $\dot{\rho}_t = \xi_\rho(\rho - \rho_t)$, where ρ denotes the steady-state level. We study a one-time, unanticipated ("MIT") shock at time $t = 0$, initializing the shock at $\rho_0 = 1.5\rho$ and calibrating its persistence to a half-life of one quarter.

Optimal monetary stabilization of demand shocks. Figure 3 plots the optimal transition dynamics under the timeless Ramsey plan in response to a demand shock.

Divine Coincidence obtains in RANK in the face of demand and productivity shocks: the planner perfectly stabilizes both the output and inflation gaps. This benchmark result requires the appropriate employment subsidy, which we assume here. To support this desired allocation, the

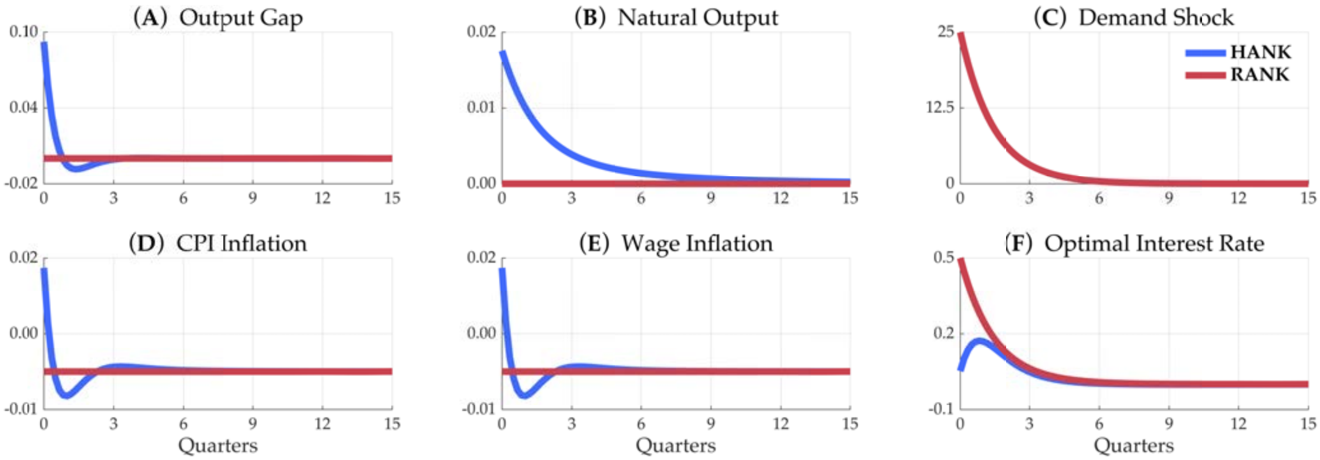


Figure 3. Optimal Policy Transition Dynamics: Demand Shock

Note. Transition dynamics after a positive discount rate shock in RANK (red) and HANK (blue) under optimal monetary stabilization policy. The discount rate shock is initialized at $\rho_0 = 0.025$ and mean-reverts to its steady state value $\rho = 0.02$, with a half-life of 1 quarter. Panels (**A**) through (**C**) report the dynamics of the output gap, $\frac{Y_t - \tilde{Y}_t}{\tilde{Y}_t}$, natural output, and the shock, all in percent deviations from the stationary Ramsey plan. Panels (**D**) through (**F**) report CPI inflation, wage inflation, and the optimal interest rate, all in percentage point deviations from the stationary Ramsey plan.

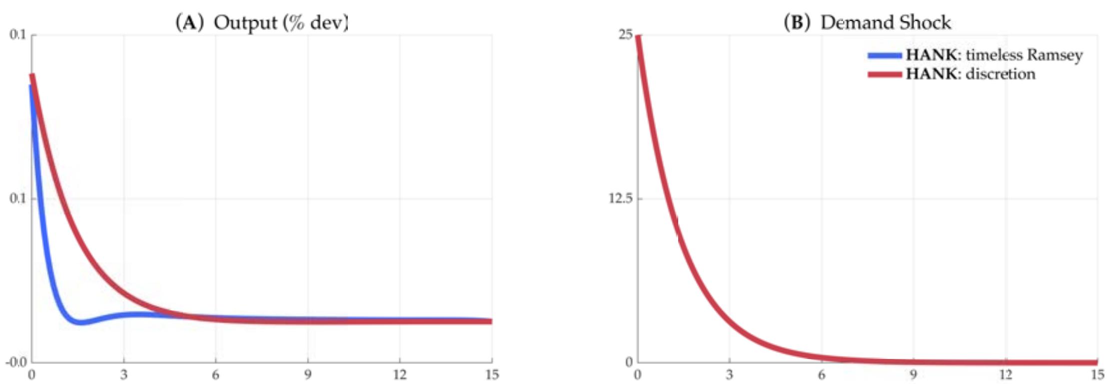


Figure 4. Optimal Policy under Discretion: Demand Shock

Note. Transition dynamics after a positive discount rate shock, comparing optimal policy in HANK with commitment (blue), i.e., under the timeless Ramsey problem, and under discretion (red). Discount rate shock is initialized at $\rho_0 = 0.025$ and mean-reverts to its steady state value $\rho = 0.02$, with a half-life of 1 quarter. Panel (**A**) plots output in percent deviation from the 0-inflation steady state for commitment (blue) and in percent deviation from the Markov perfect equilibrium with inflationary bias for discretion (red). Panel (**B**) plots the underlying shock in percent deviations.

planner raises the interest rate by about 50 basis poins to lean against the 50 basis point discount rate shock.

In HANK, the planner again leans against the demand shock, stabilizing output and inflation gaps, but not as strongly as in RANK. Especially the output gap is allowed to open up meaningfully. The on-impact output gap response under optimal policy is only dampened by 50% relative to the Taylor rule case. The inflation gap, on the other hand, is stabilized almost entirely. Unlike in RANK, the path of interest rates that supports this allocation features a hump, where the planner only gradually increases the nominal rate. The hump-shaped paths of the output and inflation gaps in Figure 3 are due to commitment—see our discussion in Section 4.6. In Figure 4, we plot the relative output paths under discretion (red) and under the timeless Ramsey problem (blue). Under discretion, the planner takes future policy as given and does not benefit from promising an over-shooting. With commitment, the planner finds it optimal to promise over-shooting to improve contemporaneous tradeoffs. This is reflected in the hump-shaped optimal interest rate path in Figure 3.

To pinpoint the source of departure from Divine Coincidence in HANK, we rely on decomposition (65). The markup wedge does not respond to a demand shock. Similarly, Figure 4 highlights that the two penalty wedges, which vanish under discretion, contribute little to the optimal on-impact response of the output gap, which is nearly identical when policy is set under discretion. Therefore, the departure from Divine Coincidence in response to a demand shock is quantitatively driven—at least on impact—by the redistribution wedge.

6 Conclusion and Broader Insights

This paper draws three main conclusions for the design of optimal monetary policy in the presence of heterogeneous households. First, a utilitarian planner under discretion trades off aggregate stabilization against a novel redistribution motive. This redistribution motive is a new source of time inconsistency that substantially exacerbates inflationary bias. In HANK, policy under discretion consequently leads to inflationary bias even with the appropriate employment subsidy to correct the markup distortion in steady state. Under commitment, the utilitarian planner recognizes that monetary policy is an inappropriate instrument to address this new source of perceived suboptimality by promising zero inflation in the long run. Second, two sets of penalties are necessary for monetary policy to be time consistent: the standard inflation penalty, which must be augmented by distributional considerations, and new distributional penalties, which penalize those individuals who benefit from discretionary policy. Abiding by these penalties allows the planner to achieve zero inflation in the long-run. The commitment solution can still be implemented by a planner under discretion, as long as she is confronted with the appropriate penalties, or through an appropriate inflation targeting framework. Third, Divine Coincidence breaks down, even in the absence of cost-push shocks, and optimal monetary stabilization policy will account for the distributional impact of policies, trading off aggregate stabilization against distributional considerations.

Three broader insights emerge from our study of optimal policy in HANK economies.

1. **HANK vs. RANK.** Household heterogeneity has stark implications for optimal monetary policy under discretion, where a new source of time inconsistency exacerbates inflationary bias. New penalties are required to make monetary policy time consistent. On the other hand, household heterogeneity in our model does not alter the optimality of 0 inflation in the long run. And while it has implications for optimal stabilization policy, departures from RANK are quantitatively small.

2. **Joint aggregate and distributional impact of policy.** Optimal policy is shaped by its joint aggregate and distributional impact. In the stylized model of this paper, lowering rates stimulates the economy and improves redistribution. However, this pattern may be different in richer environments. While our approach and the logic of our results will extend to these cases, the exact conclusions may not, which opens the door to future research. For instance, lowering interest rates may benefit wealthy, low marginal utility households, plausibly through labor market effects, credit market access, or differential inflation. Alternatively, bailouts or unconventional monetary policy may stimulate the economy but harm redistribution by favoring wealthy, low marginal utility households. Through the logic developed in this paper, a planner will have an incentive to run an underheated economy in both of these scenarios, inducing deflationary bias under discretion and requiring different inflation and distributional penalties to implement optimal policy under commitment.

3. **Role of mandates.** The commitment solution is one possible way of addressing the new source of inflationary bias identified in this paper. In RANK economies, Rogoff (1985) argues that increasing the weight on inflation in the central bank's loss function (mandate) may be valuable to reduce inflationary bias. In HANK economies, the analog to Rogoff's solution is to lower the weight on redistribution in the central bank's loss function (mandate). We study the design of central bank mandates when society values distributional considerations in ongoing and future work. In Dávila and Schaab (2022), we develop a methodology to decompose the normative considerations that determine the aggregate and redistribution consequences of welfare assessments in general economies. In Dávila and Schaab (2023), we leverage that methodology to explore the role of central bank mandates and how they impact optimal monetary under discretion and commitment.

References

Acharya, Sushant, Edouard Challe, and Keshav Dogra. 2020. Optimal monetary policy according to HANK.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. Income and wealth distribution in macroeconomics: A continuous-time approach. *The review of economic studies*, 89(1):45–86.

Ampudia, Miguel, Dimitris Georgarakos, Jiri Slacalek, Oreste Tristani, Philip Vermeulen, and Giovanni Violante. 2018. Monetary policy and household inequality.

Auclert, Adrien. 2019. Monetary policy and the redistribution channel. *American Economic Review*, 109(6):2333–2367.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2018. The intertemporal keynesian cross.

—. 2020. Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model.

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

Benigno, Pierpaolo, Gauti B Eggertsson, and Federica Romei. 2020. Dynamic debt deleveraging and optimal monetary policy. *American Economic Journal: Macroeconomics*, 12(2):310–350.

Benigno, Pierpaolo and Michael Woodford. 2012. Linear-quadratic approximation of optimal policy problems. *Journal of Economic Theory*, 147(1):1–42.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J Sargent. 2021. Inequality, Business Cycles, and Monetary-Fiscal Policy.

Bilbiie, Florin O. 2008. Limited asset markets participation, monetary policy and (inverted) aggregate demand logic. *Journal of economic theory*, 140(1):162–196.

Bilbiie, Florin Ovidiu. 2018. Monetary policy and heterogeneity: An analytical framework.

Blanchard, Olivier and Jordi Galí. 2007. Real wage rigidities and the New Keynesian model. *Journal of money, credit and banking*, 39:35–65.

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89:68–92.

Chari, Varadarajan V and Patrick J Kehoe. 1999. Optimal fiscal and monetary policy. *Handbook of macroeconomics*, 1:1671–1745.

Chari, Varadarajan V, Juan Pablo Nicolini, and Pedro Teles. 2020. Optimal capital taxation revisited. *Journal of Monetary Economics*, 116:147–165.

Clarida, Richard, Jordi Galí, and Mark Gertler. 1999. The science of monetary policy: a new Keynesian perspective. *Journal of economic literature*, 37(4):1661–1707.

Clayton, Christopher and Andreas Schaab. 2022. A Theory of Dynamic Inflation Targets.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2017. Innocent Bystanders? Monetary policy and inequality. *Journal of Monetary Economics*, 88:70–89.

Cúrdia, Vasco and Michael Woodford. 2016. Credit frictions and optimal monetary policy. *Journal of Monetary Economics*, 84:30–65.

Dávila, Eduardo and Anton Korinek. 2018. Pecuniary Externalities in Economies with Financial Frictions. *The Review of Economic Studies*, 85(1):352–395.

Dávila, Eduardo and Andreas Schaab. 2022. Welfare Assessments with Heterogeneous Individuals.

—. 2023. Central Bank Mandates with Distributional Considerations.

Doepke, Matthias and Martin Schneider. 2006. Inflation and the redistribution of nominal wealth. *Journal of Political Economy*, 114(6):1069–1097.

Erceg, Christopher J, Dale W Henderson, and Andrew T Levin. 2000. Optimal monetary policy with staggered wage and price contracts. *Journal of monetary Economics*, 46(2):281–313.

Farhi, Emmanuel and Iván Werning. 2016. A theory of macroprudential policies in the presence of nominal rigidities. *Econometrica*, 84(5):1645–1704.

Galí, Jordi. 2015. *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

González, Beatriz, Galo Nuño, Dominik Thaler, and Silvia Albrizio. 2021. Optimal monetary policy with heterogeneous firms.

Harris, Christopher and David Laibson. 2013. Instantaneous gratification. *The Quarterly Journal of Economics*, 128(1):205–248.

Kamien, Morton I and Nancy Lou Schwartz. 2012. *Dynamic optimization: the calculus of variations and optimal control in economics and management*. courier corporation.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. Monetary policy according to HANK. *American Economic Review*, 108(3):697–743.

Khan, Aubhik, Robert G King, and Alexander L Wolman. 2003. Optimal monetary policy. *The Review of Economic Studies*, 70(4):825–860.

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

—. 1980. Dynamic optimal taxation, rational expectations and optimal control. *Journal of Economic Dynamics and control*, 2:79–91.

Le Grand, François, Alaïs Martin-Baillon, and Xavier Ragot. 2021. Should monetary policy care about redistribution? Optimal fiscal and monetary policy with heterogeneous agents.

Ljungqvist, Lars and Thomas J Sargent. 2018. *Recursive macroeconomic theory*. MIT press.

Marcet, Albert and Ramon Marimon. 2019. Recursive contracts. *Econometrica*, 87(5):1589–1631.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158.

McKay, Alisdair and Christian K Wolf. 2022. Optimal Policy Rules in HANK.

Nuño, Galo and Benjamin Moll. 2018. Social optima in economies with heterogeneous agents. *Review of Economic Dynamics*, 28:150–180.

Nuño, Galo and Carlos Thomas. 2020. Optimal monetary policy with heterogeneous agents.

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

Rotemberg, Julio J. 1982. Sticky prices in the United States. *Journal of Political Economy*, 90(6):1187–1211.

Sannikov, Yuliy. 2008. A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984.

Schaab, Andreas. 2020. Micro and Macro Uncertainty.

Schaab, Andreas and Allen Tianlun Zhang. 2022. Dynamic Programming in Continuous Time with Adaptive Sparse Grids.

Schaumburg, Ernst and Andrea Tambalotti. 2007. An investigation of the gains from commitment in monetary policy. *Journal of monetary economics*, 54(2):302–324.

Schmitt-Grohé, Stephanie and Martín Uribe. 2004a. Optimal fiscal and monetary policy under sticky prices. *Journal of economic Theory*, 114(2):198–230.

—. 2004b. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775.

—. 2005. Optimal fiscal and monetary policy in a medium-scale macroeconomic model. *NBER Macroeconomics Annual*, 20:383–425.

—. 2007. Optimal simple and implementable monetary and fiscal rules. *Journal of monetary Economics*, 54(6):1702–1725.

—. 2010. The optimal rate of inflation. In *Handbook of monetary economics*, volume 3, pp. 653–722.

Smirnov, Danila. 2022. Optimal Monetary Policy in HANK.

Svensson, Lars E O. 1997. Optimal Inflation Targets, "Conservative" Central Banks, and Linear Inflation Contracts. *The American Economic Review*, 87(1):98–114.

Werning, Ivan. 2011. Managing a liquidity trap: Monetary and fiscal policy.

Williams, Noah. 2011. Persistent private information. *Econometrica*, 79(4):1233–1275.

Woodford, Michael. 1999. Commentary: How should monetary policy be conducted in an era of price stability? *New challenges for monetary policy*, 277316.

—. 2003. *Interest and prices*. Princeton University Press.

—. 2010. Optimal monetary stabilization policy. *Handbook of monetary economics*, 3:723–828.