CESifo AREA CONFERENCES 2022

Macro, Money, and International Finance
21 – 22 July 2022

Optimal Monetary Policy with Heterogeneous Agents: A Timeless Ramsey Approach

Eduardo Dávila and Andreas Schaab

CESifo

Optimal Monetary Policy with Heterogeneous Agents: A Timeless Ramsey Approach

Eduardo Dávila[*] Andreas Schaab[†]

First version: March, 2022
This version: July, 2022

Abstract

This paper characterizes optimal monetary policy in a canonical heterogeneous-agent New Keynesian ("HANK") model with wage rigidity. After characterizing optimal policy under discretion, we develop a timeless Ramsey approach to study i) optimal long-run policy, ii) time inconsistency and targeting rules, and iii) optimal stabilization policy. We show that zero inflation is the optimal long-run policy under commitment, while policy under discretion leads to a inflationary bias exacerbated by a novel redistribution motive: the planner has a time-inconsistent incentive to lower the real interest rate—thus overheating the economy—to redistribute towards indebted, high marginal utility households. Time-consistent monetary policy requires both a standard inflation target that accounts for distributional considerations as well as a novel distributional target. Optimal stabilization policy trades off aggregate stabilization against redistribution: Divine Coincidence fails in heterogeneous-agent economies even in the absence of cost-push shocks due to the distributive pecuniary effects of interest rate policy. Finally, we show that there are gains from commitment even in the absence of cost-push shocks. We compute optimal stabilization policy in response to productivity, demand, and cost-push shocks both non-linearly and using sequence-space perturbation methods, which we extend to Ramsey problems and welfare analysis.

JEL codes: E52, E61

Keywords: optimal monetary policy, heterogeneous-agent New Keynesian model, policy under discretion, timeless Ramsey approach, timeless penalties, inflation target, distributional target, sequence-space methods

[*]Yale University and NBER. Email: eduardo.davila@yale.edu. Website: https://eduardodavila.com.

[†]Columbia Business School. Email: ajs2428@columbia.edu. Website: https://andreasschaab.com.

We are grateful to Mark Aguiar, Manuel Amador, Marios Angeletos, Larry Christiano, Marty Eichenbaum, Emmanuel Farhi, Mike Golosov, Greg Kaplan, Benjamin Moll, Emi Nakamura, Galo Nuño, Matt Rognlie, and Ludwig Straub for helpful conversations, as well as seminar participants at Columbia, Northwestern, and the College de France Conference in Honor of Emmanuel Farhi.

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., 2020). Household heterogeneity may therefore be an important determinant of the welfare impact of monetary policy and should inform the study of optimal policy design.[1] However, accounting for rich heterogeneity and incomplete markets in dynamic optimal policy problems has remained challenging because the planner must internalize the effects of policy on an evolving cross-sectional distribution.

In this paper, we develop a *timeless Ramsey approach* to study optimal monetary policy in heterogeneous-agent environments. This approach allows us to isolate three important dimensions of monetary policy design: i) long-run policy, ii) time consistency and targeting rules, and iii) stabilization policy. Household heterogeneity may interact with optimal policy considerations along each of these dimensions. We leverage the timeless Ramsey approach to systematically revisit the canonical New Keynesian consensus on optimal monetary policy (Clarida et al., 1999; Woodford, 2003; Galí, 2015) and study—both analytically and quantitatively—the implications of household heterogeneity.

We introduce our approach and develop our analytical results in a one-asset HANK economy with wage rigidity, which represents a minimal departure from the representative-agent New Keynesian (RANK) model. Our analysis of optimal monetary policy is structured parallel to 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.

Policy under discretion. We initially study policy under discretion for two reasons. First, the economic rationales that determine optimal monetary policy with heterogeneity emerge more clearly in the problem under discretion. In fact, allowing for commitment simply introduces promise-keeping considerations. Second, understanding policy under discretion motivates our subsequent analysis of commitment and targeting rules. In particular, we show that a utilitarian planner's incentive to redistribute to indebted, high marginal utility households by lowering interest rates exacerbates the classical inflationary bias problem, thus implying even larger benefits

[1] In fact, the Federal Reserve has started taking into account "distributional considerations." After concluding a long-run strategic review in August, 2020, Chairman Jerome Powell remarked: "*Our revised statement emphasizes that maximum employment is a broad-based and inclusive goal. This change reflects our appreciation for the benefits of a strong labor market, particularly for many in low- and moderate-income communities.*" The full speech can be found at: `https://www.federalreserve.gov/newsevents/speech/powell20200827a.htm`.

from commitment and credibility

An important and well-understood insight from representative-agent analysis is that there are no gains from commitment in the absence of cost-push shocks when the planner sets an appropriate employment subsidy in steady state. No time consistency problem emerges in this case, and there are no gains from commitment relative to discretion. This important benchmark result fails in an environment with heterogeneous agents.

In a HANK economy, a utilitarian planner under discretion has an incentive to raise output above natural output and overheat the economy, even in the absence of markup distortions. The fact that the distributive pecuniary effect associated with lowering interest rates favors indebted, high marginal utility households provides a rational for the discretionary planner to depress the real interest rate. This new redistribution motive incentivizes a utilitarian planner to overheat the economy, generating surprise inflation ex-post, regardless of whether markup distortions are present. When agents anticipate the planner's incentives, an inflationary bias emerges in equilibrium. Hence, the novel redistribution motive exacerbates the inflationary bias problem, increasing the gains from commitment. Quantitatively, we find that the inflationary bias generated by the redistribution motive is about 4 times larger than the classical source of inflationary bias stemming from the Phillips curve (Kydland and Prescott, 1977; Barro and Gordon, 1983).

Motivated by these results, we develop a timeless Ramsey approach to study optimal policy under commitment in Section 4. The timeless Ramsey approach comprises three steps. Each step allows us to isolate one important dimension of monetary policy design: long-run policy, time consistency and targeting rules, and stabilization policy. Our goal is to study the implications of household heterogeneity for optimal policy along each of these dimensions.

Optimal long-run policy. In the first step of our timeless Ramsey approach, we define a standard Ramsey problem in primal form and characterize its solution, which we refer to as a Ramsey plan.[2] The primal Ramsey problem is constrained by the implementability conditions that characterize competitive equilibrium. At this stage, Ramsey plans conflate three distinct economic motives that govern optimal policy dynamics—incentives to respond to long-run distortions, time consistency problems, and stabilization motives in response to shocks. We first solve for the stationary equilibrium that the Ramsey planner perceives as optimal in the long run—a stationary Ramsey plan—and characterize the policy that supports it—a stationary Ramsey policy. Crucially, incentives to respond to transient shocks or time consistency problems dissipate in the long run. Hence, characterizing the stationary Ramsey plan allows us to isolate which policy considerations shape the optimal long-run policy.

We show that the optimal long-run inflation rate in our baseline HANK economy is zero.

[2] 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. Alternatively, 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).

Therefore, monetary policy under commitment does not suffer from the inflationary bias problem that plagues policy under discretion in the long run. The RANK limit of our economy also features zero optimal long-run inflation. Consequently, household heterogeneity has no implications for optimal long-run policy in our baseline environment. This result critically hinges on the fact that the nominal interest rate and inflation enter the Ramsey problem symmetrically, which can be seen as a relevant benchmark. That is, in environments in which the nominal interest and inflation have a differential impact on different individuals in the economy, we should expect a stationary Ramsey policy that features non-zero inflation in the long run, even without other considerations that are known to affect the optimal long-run rate of inflation, e.g., a demand for fiat money.

Time consistency and targeting rules. In the second step of our timeless Ramsey approach, we show that standard Ramsey plans feature two dimensions of time inconsistency. The first source of time inconsistency arises because expectations about future inflation enter the Ramsey problem via the forward-looking New Keynesian wage Phillips curve. 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 different source of time consistency—entirely absent from RANK economies—emerges. As in the problem under discretion, the Ramsey planner always has an incentive to lower interest rates, redistributing towards indebted, high marginal utility households. To overcome such a desire to redistribute, the Ramsey planner must put in place promise-keeping constraints based on households' lifetime utilities. That is, the way in which a planner commits to not lowering interest rates to redistribute is by penalizing welfare gains by indebted, high marginal households when making welfare assessments.[3] These penalties are engineered to counteract the planner's desire to redistribute. Time inconsistency is simply a manifestation of the planner's desire to break the promise-keeping constraints or, equivalently, to ignore the penalties.[4]

While standard Ramsey plans address the problem of inflationary bias in the long run by building up promises, they still suffer from inflationary bias in the short run. Both time consistency problems provide an incentive for the planner to overheat the economy in the short run. Motivated by this observation, we introduce *timeless penalties* for each of the forward-looking implementability conditions, extending the approach of Marcet and Marimon (2019) to our setting (i.e., continuous-time heterogeneous-agent economies). We define a *timeless Ramsey problem* that augments the standard Ramsey problem with these timeless penalties and show that it yields time-consistent planning solutions. That is, the timeless Ramsey planner no longer faces an incentive to overheat the economy either in the long run or in the short run.

[3] Formally, the implementability conditions of the Ramsey problem include a cross section of forward-looking equations that depend on future interest rates and that characterize individual consumption and savings decisions. By judiciously designing promise-keeping constraints, the planner can credibly commit to a future paths of interest rates.

[4] By definition, a standard Ramsey planner faces no promise-keeping constraints at date 0, and builds promises over time. Hence, time-inconsistent concerns are maximal at the stationary Ramsey plan.

We show that the timeless penalties required for time-consistent policy in HANK can be interpreted as a combination of an inflation target and a novel distributional target. Connecting to the vast RANK literature on time consistency and targeting rules, we establish three main results.

First, the inflation target in HANK economies takes the standard form of a linear penalty on inflation, which is now also shaped by distributional considerations. In the presence of household heterogeneity, the utilitarian planner's redistribution motive interacts with the classical time consistency problem on inflation that emerges from the Phillips curve. The augmented inflation target that supports time-consistent monetary policy in HANK reflects these interactions. In particular, the benchmark result of the representative-agent literature is that no inflation target is necessary in the absence of cost-push shocks when the planner can set an appropriate employment subsidy to offset markup distortions in steady state. No time consistency problem emerges in this case. This important benchmark result breaks down in heterogeneous-agent environments, because a utilitarian planner has an incentive to use inflation to depress the real interest rate to benefit indebted, high marginal utility households. Whenever a planner (central bank) values distributional considerations, the inflation target necessary to conduct time-consistent policy must account for this distributional motive.

Second, since monetary policy faces a second source of time inconsistency in HANK, a planner (central bank) that adopts a welfare criterion (mandate) that is not the aggregate efficiency one must also adopt a new *distributional target* in addition to the standard inflation target in order to implement time-consistent policy. Similar to the inflation target, this distributional target takes the form of a linear penalty on lifetime utilities. In particular, it confronts the planner with a penalty for redistributing towards indebted, high marginal utility households. Formally, we show that the distribution of promises associated with this form of time inconsistency evolves according to a particular Kolmogorov forward equation. We leverage this equation to characterize the distributional target that a planner must adopt to conduct time-consistent monetary policy.

Third, we justify the interpretation of our timeless penalties as targets by showing that policy under discretion no longer features inflationary bias in steady state when confronted with the augmented inflation target and the novel distributional target. We relate our results to the literature on transferable utility mechanisms and point out the resemblance between the timeless penalties we derive and the optimal mechanisms that typically emerge in that literature.

Optimal stabilization policy. In the third and final step of our timeless Ramsey approach, we study optimal stabilization policy under the timeless Ramsey plan. Under the appropriate timeless penalties, optimal policy dynamics in response to shocks are now driven solely by stabilization motives and are no longer confounded by considerations of long-run distortions and time inconsistency. We characterize an analytical targeting rule for optimal monetary stabilization policy under commitment that illustrates the departures from optimal policy in RANK, which it nests.

In a RANK economy, the Divine Coincidence result establishes that no tradeoff emerges between inflation and output in the absence of cost-push shocks; the planner finds it optimal to close both the inflation and output gaps at the same time (Blanchard and Galí, 2007). In HANK economies, on the other hand, Divine Coincidence generically fails even in the absence of cost-push shocks. Due to the distributive pecuniary effects of interest rate policy, the planner always perceives a tradeoff between aggregate stabilization, i.e., inflation and output, on the one hand, and distributional considerations on the other hand. Accounting for such distributional considerations comes at the cost of aggregate efficiency. Indeed, from the perspective of a planner (central bank) that only values aggregate efficiency, the Divine Coincidence benchmark is restored in our setting.

We provide a quantitative analysis of optimal stabilization policy in Section 5, where we leverage our timeless Ramsey approach to compute optimal policy dynamics in response to three types of shocks—demand, productivity, and cost-push shocks. While our approach allows us to characterize and compute timeless Ramsey plans non-linearly, an important contribution of this paper is to bring perturbation methods to bear on the question of optimal stabilization policy in heterogeneous-agent economies. Building on the increasingly popular literature on sequence-space methods, we develop a sequence-space representation of timeless Ramsey plans. Under this representation, a Ramsey plan is a system of equations that takes as inputs the time paths of aggregate allocations and prices, aggregate multipliers, policies, and shocks. While we can also leverage this representation to compute Ramsey plans non-linearly, we develop sequence-space perturbation methods—in both a primal and dual form—to approximate optimal stabilization policy to first order. 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 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. Our paper therefore builds on Boppart et al. (2018) and Auclert et al. (2021) and extends the sequence-space apparatus to Ramsey problems and welfare analysis in heterogeneous-agent environments.

Related literature. Our paper contributes to multiple branches of the literature on optimal monetary policy. First and foremost, our paper contributes to the growing literature on optimal policy in HANK economies. This literature includes the work of Bhandari et al. (2021), who introduce a small-noise expansion method to compute optimal monetary and fiscal policy in HANK models; Acharya et al. (2020), who 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), who study optimal monetary and fiscal policy keeping heterogeneity finite-dimensional by truncating idiosyncratic histories; González et al. (2021), who study optimal monetary policy in an environment with firm heterogeneity; and McKay and Wolf (2022), who study optimal monetary policy with heterogeneous households in linear-quadratic environments.[5]

Within this literature, our continuous-time formulation of the optimal monetary policy problem, treating the cross-sectional distribution of households as a control, is most closely related to the work of Nuño and Thomas (2020), on which we build.[6] Nuño and Thomas (2020) study optimal monetary policy in a small open economy, in which short-term real interest rates and output are unaffected by monetary policy.[7] A contribution of our paper is to study optimal monetary policy in a closed economy that features the classic output-inflation tradeoff, which is central to the New Keynesian literature and allows us to characterize a new source of time inconsistency due to the presence of individual heterogeneity. As in their paper, we characterize fully dynamic, non-linear Ramsey plans. We also develop sequence-space perturbation methods to compute optimal stabilization policy around the stationary Ramsey plan.

Both in the body of the paper and in the Appendix, we purposefully relate our results to those of the vast literature on monetary policy in RANK models (Clarida et al., 1999; Woodford, 2003; Galí, 2015). By doing so we are able to provide clear analytical insights into the form of the optimal policy in a HANK environment and how it relates to the well-understood results in RANK economies. At an abstract level, our approach is closest to the work of Khan et al. (2003), who initially characterize standard and timeless Ramsey optimal policy using the exact structural equations and utility function, and then use perturbation methods to characterize the optimal responses to shocks. Schmitt-Grohé and Uribe (2010) and Woodford (2010) systematically study and review optimal long-run policy and optimal stabilization policy in RANK economies.

Our characterization of timeless penalties builds on the recursive multiplier approach of Marcet and Marimon (2019). We provide a novel application of their approach in a continuous-time environment where the planner must keep track of a full cross-sectional distribution and faces a continuum of individual forward-looking constraints. A central contribution of our paper is to show that the evolution of the distribution of promises associated with such forward-looking constraints under the optimal Ramsey plan satisfies a Kolmogorov forward equation that accounts for the "births" and "deaths" of promises.

Finally, we contribute to recent work on computational methods in heterogeneous-agent environments by extending the sequence-space apparatus to Ramsey problems and welfare analysis (Boppart et al., 2018; Auclert et al., 2021). We develop a sequence-space representation of timeless Ramsey plans, which we can compute non-linearly or 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 form. We also introduce and define sequence-space Hessians as the natural, second-order generalization of sequence-space Jacobians.

[5] Benigno et al. (2020) study optimal monetary policy in a two-agent New Keynesian model.

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

[7] 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 (32)—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.

2 Baseline Model

Our baseline model is a one-asset heterogeneous-agent New Keynesian ("HANK") model with wage rigidity. This model represents a minimal departure from a representative-agent New Keynesian ("RANK") model (Clarida et al., 1999; Woodford, 2003; Galí, 2015), which we use as a benchmark for our analytical and quantitative results. Our baseline model is deliberately stylized to make our characterization of optimal monetary policy accessible.

We cast our baseline model in continuous time. The time horizon is infinite with $t \in [0, \infty)$. There is no aggregate uncertainty and we focus on one-time, unanticipated shocks. In particular, following much of the New Keynesian literature, we allow for three types of shocks: demand, productivity, and cost-push shocks.

2.1 Households

The economy is populated by a unit mass of households who consume and work. Household preferences are given by

$$\mathbb{E}_0 \int_0^\infty e^{-\int_0^t \rho_s ds} \left[u(c_t) - \Phi\left(\left\{ n_{k,t}, \pi^w_{k,t} \right\}_{k \in [0,1]} \right) \right] dt, \tag{1}$$

where c_t is the rate of consumption, $u(\cdot)$ captures the instantaneous utility flow from consumption, and ρ_t is a common but potentially time-varying discount rate, which represents a source of demand shocks. The function $\Phi(\cdot)$ captures the household's disutility from work and its details depend on the wage bargaining and labor market structure, which we discuss below. In particular, $n_{k,t}$ denotes the hours that the household supplies to union k and $\pi^w_{k,t}$ is union k's wage inflation, which enters $\Phi(\cdot)$ as an additional cost, as we describe below.

Each household supplies labor to all of $k \in [0, 1]$ unions. We denote a household's total hours of work by

$$n_t = \int_0^1 n_{k,t} dk.$$

Each union pays the household a nominal wage $W_{k,t}$. The household budget constraint therefore corresponds to

$$\dot{a}_t = r_t a_t + z_t \frac{1}{P_t} \int_0^1 W_{k,t} n_{k,t} dk + \tau_t(z_t) - c_t, \tag{2}$$

where a_t denotes bond holdings, r_t the real interest rate, and P_t is the price of the consumption good. Households' non-financial income comprises labor income, which is proportional to idiosyncratic labor productivity z_t, and lump-sum rebates $\tau_t(z_t)$, which may also depend on z_t. Finally, households face a borrowing constraint given by

$$a_t \geq \underline{a},$$

where $\underline{a} \leq 0$.

Even though there is no aggregate uncertainty, households face idiosyncratic earnings risk. We assume that labor productivity z_t follows an exogenous Markov process that we further specialize below. Since households are only heterogeneous ex-post, we can index individual households by their idiosyncratic state variables (a, z). We denote the mass of households with idiosyncratic states (a, z) by $g_t(a, z)$, which we also refer to as the cross-sectional distribution. And since our economy is populated by a measure 1 of infinitely-lived households, we have $\iint g_t(a, z)\, da\, dz = 1$.

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., 2020), labor unions determine work hours. Each union $k \in [0, 1]$ transforms hours supplied by households into a differentiated labor service according to the linear aggregation technology

$$N_{k,t} = \iint z n_{k,t} g_t(a, z)\, da\, dz,$$

where $N_{k,t}$ is expressed in units of effective labor. Each union also rations labor, so that all households work the same hours. In particular, this implies $N_{k,t} = n_{k,t} \iint z g_t(a, z)\, da\, dz = n_{k,t}$, after normalizing cross-sectional average labor productivity to 1.

Labor packer. Unions sell their differentiated labor services to an aggregate labor packer. The packer operates the CES aggregation technology

$$N_t = \left(\int_0^1 N_{k,t}^{\frac{\epsilon_t - 1}{\epsilon_t}}\, dk \right)^{\frac{\epsilon_t}{\epsilon_t - 1}},$$

where the elasticity of substitution ϵ_t is potentially time-varying. We interpret time variation in the desired wage mark-up of unions as a source of cost-push shocks, following standard practice (see, e.g., Galí, 2015). The packer sells the aggregate labor bundle to firms at nominal wage rate W_t. The labor packer's cost-minimization problem is standard and yields the demand function and wage index

$$N_{k,t} = \left(\frac{W_{k,t}}{W_t} \right)^{-\epsilon_t} N_t \tag{3}$$

$$W_t = \left(\int_0^1 W_{k,t}^{1-\epsilon_t}\, dk \right)^{\frac{1}{1-\epsilon_t}}, \tag{4}$$

where $W_{k,t}$ is the nominal wage rate charged by union k.

Wage rigidity. Nominal wages are sticky in our model. Each union k faces an adjustment cost to change its wage. Formally, the union takes $W_{k,t}$ as a state variable and controls how the wage evolves by setting wage inflation $\pi^w_{k,t}$, with

$$\pi^w_{k,t} = \frac{\dot{W}_{k,t}}{W_{k,t}}. \tag{5}$$

In our baseline model, the union's adjustment cost is directly passed to union members as a quadratic utility cost, so $\Phi(\cdot)$ in equation (1) is explicitly given by

$$\Phi\left(\left\{n_{k,t}, \pi^w_{k,t}\right\}_{k\in[0,1]}\right) = v\left(\int_0^1 n_{k,t}\, dk\right) + \frac{\delta}{2}\int_0^1 (\pi^w_{k,t})^2 dk,$$

where $v(\cdot)$ captures pure disutility from working and δ modulates the strength of the wage rigidity.

Employment subsidy. As is standard in the New Keynesian literature, we allow for an employment subsidy. Given union wage receipts $z_t W_{k,t} n_{k,t}$ to a household with labor productivity z_t, the government pays the household a proportional income subsidy $\tau^L z_t W_{k,t} n_{k,t}$, which the union internalizes when setting wages.

Wage Phillips curve. We now formalize the union's wage setting problem to derive a New Keynesian wage Phillips curve. We assume that the union chooses wages in order to maximize stakeholder value—the sum of stakeholders', i.e., union members', utilities.[8] That is, union k solves

$$\max_{\pi^w_{k,t}} \int_0^\infty e^{-\int_0^t \rho_s ds}\left(\iint \left[u(c_t(a,z;W_{k,t})) - v\left(\int_0^1 n_{k,t}\, dk\right) - \frac{\delta}{2}\int_0^1 (\pi^w_{k,t})^2\, dk\right] g_t(a,z)\, da\, dz\right) dt \tag{6}$$

subject to equations (3) and (5). The union further internalizes the effect of its wage policy on its members' consumption—hence the explicit dependence of c_t on $W_{k,t}$ in equation (6). However, since union k is small, it takes as given all macroeconomic aggregates, including the cross-sectional household distribution. We solve the union's problem in Section C of the Appendix, where we also derive wage Phillips curves under alternative assumptions on wage adjustment costs.[9]

There, we show that the wage policies that result from the union's problem give rise to a symmetric equilibrium, where wages and labor allocations are equalized across unions, i.e., $W_{k,t} = W_t$ and $N_{k,t} = N_t$. In such a symmetric equilibrium, the non-linear New Keynesian wage

[8] The choice of objective by a union under incomplete markets is subject to the same caveats as the choice of objective by a firm. It is straightforward to consider alternative objectives.

[9] There are three natural ways to model wage adjustment costs: as an explicit resource cost that is passed on to households, as labor productivity distortions, or as a direct utility cost. In the main text, we adopt the utility cost specification largely because it is most tractable. For robustness, in the Appendix, we derive alternative Phillips curves under different assumptions, and discuss how they impact our conclusions.

Phillips curve is given by

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta}\left[\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) w_t \iint z u'(c_t(a,z)) g_t(a,z)\, da\, dz - v'(N_t)\right] N_t,$$

highlighting that demand and cost-push shocks directly affect wage inflation dynamics.

Labor wedges. Labor wedges will play a key role in our analysis of welfare and optimal policy. We introduce them now and show that they provide an insightful interpretation of the wage Phillips curve. Following Farhi and Werning (2016), we define the *individual labor wedge* as $\tilde{\tau}_t(a,z) = w_t z u'(c_t) - v'(N_t)$. We also define the *desired-markup augmented individual labor wedge*, which we simply refer to as the augmented labor wedge, as

$$\tau_t(a,z) = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) w_t z u'(c_t) - v'(N_t)\right) N_t. \tag{7}$$

We can now reformulate the New Keynesian wage Phillips curve as

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta} \iint \tau_t(a,z) g_t(a,z)\, da\, dz, \tag{8}$$

which highlights that unions' target when adjusting nominal wages is to set the aggregate augmented labor wedge to 0.[10]

2.3 Final Good Producer

A representative firm produces the final consumption good, operating the linear production technology

$$Y_t = A_t N_t, \tag{9}$$

where the aggregate labor bundle is the only input to production. We refer to A_t as total factor productivity (TFP), which is potentially time-varying and represents a source of productivity shocks. Under perfect competition and flexible prices, the real marginal cost of labor is equal to its marginal product, with

$$w_t = A_t \tag{10}$$

where $w_t = \frac{W_t}{P_t}$ denotes the real wage. Moreover, profits are zero, which is consistent with the absence of profits in equation (2).

[10] Consider the limit where wages are perfectly flexible, i.e., $\delta \to 0$, there are no cost-push shocks, and an appropriate employment subsidy ensures that $\frac{\epsilon-1}{\epsilon}(1+\tau^L) = 1$. Then, unions' target is a 0 aggregate labor wedge, which implies that wages are equalized with the aggregate marginal rate of substitution between consumption and hours, $w_t = \iint \frac{v'(N_t)}{zu'(c_t)} g_t(a,z)\, da\, dz$.

2.4 Government

We keep the role of the fiscal authority deliberately minimal in our baseline model. Our focus is on the monetary authority, which optimally sets the nominal interest rate.

Fiscal policy. There is no government spending and no debt, with bonds in zero net supply. The fiscal authority pays an employment subsidy to households. Running a balanced budget, it pays for these outlays with a lump-sum tax based on aggregate employment. We assume that both the subsidy and the tax are proportional to a household's labor productivity. That is, the net fiscal rebate that a household with idiosyncratic labor productivity z receives is zero, with

$$P_t \tau_t(z) = \int_0^1 \tau^L z W_{k,t} n_{k,t} dk - \tau^L z W_t N_t = 0.$$

Given this form of fiscal policy and the structure of the labor market, we can simplify and rewrite the household budget constraint as

$$\dot{a}_t = r_t a_t + z_t w_t N_t - c_t.$$

Monetary policy. The central bank sets the nominal interest rate i_t optimally as we describe in the next section. A Fisher relation holds in our economy, with

$$r_t = i_t - \pi_t, \tag{11}$$

where π_t is consumer price index (CPI) inflation. Finally, we can relate price inflation to wage inflation by differentiating equation (10), which yields

$$\pi_t = \pi_t^w - \frac{\dot{A}_t}{A_t}. \tag{12}$$

2.5 Equilibrium

Definition 1. (**Competitive Equilibrium**) *Given an initial distribution over household bond holdings and idiosyncratic labor productivities, $g_0(a, z)$, a symmetric initial nominal wage distribution, $W_{k,0} = W_0$, 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 for the joint distribution over household bond holdings and idiosyncratic labor productivities $\{g_t(a, z)\}$, such* *that households optimize, unions optimize, final good producers optimize, and markets clear, that is,*

$$Y_t = C_t = \iint c_t(a,z)\, g_t(a,z)\, da\, dz \tag{13}$$

$$0 = B_t = \iint a\, g_t(a,z)\, da\, dz. \tag{14}$$

In Lemma 13 in the Appendix, we provide a complete description of each of the equations that characterize an equilibrium and formally describe the implementability conditions that act as constraints for a Ramsey planner. It is helpful to separate the constraints into two blocks: an individual or "micro" block of three equations, which characterizes household consumption and utility, as well as the law of motion of the distribution of households, and a "macro block", of two equations, which characterizes the economy's aggregate resource constraint and the evolution of inflation. While the macro block is almost identical to the RANK version of our model, as we explain below, the individual block is richer.

Micro block. The individual block consists of three equations. First, there is the households' Hamilton-Jacobi-Bellman equation, given by

$$\rho_t V_t(a,z) = \partial_t V_t(a,z) + u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 + \mathcal{A}_t V_t(a,z), \tag{15}$$

where $V_t(a,z)$ denotes the lifetime utility of a household with bond holdings and idiosyncratic labor productivity (a,z), and where $\mathcal{A}_t$ denotes the infinitesimal generator of the process (a_t, z_t), formally defined in equation (53) in Appendix A.[11] Second, there is the savings optimality condition that relates the household's consumption policy function to the value function, and is given by

$$u'(c_t(a,z)) = \partial_a V_t(a,z). \tag{16}$$

Finally, there is the Kolmogorov forward equation (KFE), which describes the time evolution of the cross-sectional distribution of households over the possible individual states, (a,z), given by

$$\partial_t g_t(a,z) = \mathcal{A}_t^* g_t(a,z), \tag{17}$$

where $\mathcal{A}_t^*$ denotes the adjoint of $\mathcal{A}_t$.[12] Intuitively, the operator $\mathcal{A}_t$ accounts for the fact that households idiosyncratic states vary over time. In particular, it captures how a household's value function accounts for the path of future idiosyncratic productivity shocks and bond holdings. In turn, the adjoint $\mathcal{A}_t^*$ keeps track of how the distribution of households over the set of idiosyncratic states varies over time. Hence, it is natural for $\mathcal{A}_t^*$ to be the adjoint of $\mathcal{A}_t$. Both $\mathcal{A}_t$ and $\mathcal{A}_t^*$ are

[11] We adopt the more compact notation $\partial_x f$ to represent $\frac{\partial f}{\partial x}$.

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

Macro block. The macro block comprises the aggregate resource constraint of the economy, which we obtain by combining the goods market clearing condition (13) with the aggregate production function (9),

$$\iint c_t(a,z)\, g_t(a,z)\, da\, dz = A_t N_t, \tag{18}$$

where we normalize $\iint z g_t(a,z)\, da\, dz = 1$, and the New Keynesian wage Phillips curve (8).

2.6 Sources of Suboptimality

Before characterizing optimal policy, it is worth describing the sources of suboptimality in our baseline HANK model with wage rigidity. This economy features four sources of inefficiency, whose implications shape different dimensions of optimal policy, as we substantiate below.[13]

First, the model features monopolistic competition. Labor unions are monopolistically competitive and charge a wage markup relative to the perceived (utility) marginal cost of work hours by members. This wage markup drives a wedge between the marginal rate of transformation (MRT), A_t, and the economy's average marginal rate of substitution (MRS), Λ_t. When we set the employment subsidy so that $\frac{\epsilon-1}{\epsilon}(1+\tau^L) = 1$, the stationary equilibrium with constant elasticity of substitution ϵ features no wage markups.

Second, the model has a nominal rigidity. Nominal wages are sticky and can only be adjusted gradually in our model. This imposes two separate costs from a welfare perspective. First, assuming a steady state employment subsidy, the economy's average MRS converges only gradually to the MRT in response to shocks. Second, wage adjustments are directly associated with a deadweight (utility) cost.

Third, our model features labor rationing. We assume that unions ration work hours across their members, imposing that all households supply the same number of hours. While an appropriate notion of *average* MRS is equal to the MRT in our economy whenever the employment subsidy is in place, individual MRS are not equalized across households.

Finally, there are incomplete markets for risk. Noncontingent bonds are the only financial asset in this economy. Households furthermore face a borrowing constraint and are consequently not able to fully insure against idiosyncratic earnings risk. As a result, households' marginal rates of substitution are not equalized across periods and states.

The first two sources of inefficiency are also present in the representative-agent version of our economy. In fact, these two distortions exactly mirror the sources of inefficiency in the standard RANK model (Clarida et al., 1999; Galí, 2015). The latter two sources of inefficiency are unique to the heterogeneous-agent environment. Due to these two inefficiencies, the flexible-price allocation with an employment subsidy is no longer first-best in a HANK economy, as we show below. Consequently, the celebrated Divine Coincidence result (Blanchard and Galí, 2007) fails in our

[13] 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.

environment, as we show in Section 4.5.

2.7 Comparison Benchmarks

In the analysis that follows, we make reference to two comparison benchmarks.

The RANK limit. In order to relate to the vast literature on optimal monetary policy in the standard New Keynesian model and to better pinpoint how individual heterogeneity affects the conduct of optimal policy, it is helpful to characterize the representative-agent limit of our model as a benchmark for comparison. Starting from the baseline model, the RANK limit obtains when *i*) households' idiosyncratic labor productivity converges to a constant value, that is, $z_t \to \bar{z}$ for all t, and *ii*) the economy is initialized with a cross-sectional distribution of bond holdings and productivities that is degenerate at $a = 0$ and $z = \bar{z}$, that is, $g_0^{\text{RA}}(a,z) = \delta(a = 0, z = \bar{z})$, where $\delta(\cdot)$ denotes a two-dimensional Dirac delta function. In Appendix D, we include a self-contained characterization of the competitive equilibrium and optimal policy in the RANK limit of our economy.[14] We use $^{\text{RA}}$ superscripts to denote the RANK limit of our economy.

The flexprice allocation and natural output. As in the canonical analysis of the standard New Keynesian model, we define the flexprice allocation as the limit of our economy as $\delta \to 0$. We refer to *natural output* as the output that obtains in the flexprice allocation, denoted $\tilde{Y}_t$.[15] In the following, we sometimes use isoelastic (CRRA) preferences, with $u(c) = \frac{1}{1-\gamma}c^{1-\gamma}$ and $v(n) = \frac{1}{1+\eta}n^{1+\eta}$, to illustrate our analytical results. With isoelastic preferences, natural output is given by

$$\tilde{Y}_t = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)A_t^{1+\eta} \iint \frac{zu'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz \right)^{\frac{1}{\gamma+\eta}}, \tag{19}$$

where the integral term reflects labor rationing. In the RANK limit, where this integral term vanishes, natural output is simply given by $\tilde{Y}_t^{\text{RA}} = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)A_t^{1+\eta}\right)^{\frac{1}{\gamma+\eta}}$.

[14] The RANK limit of our economy can be summarized by two non-linear ODEs, the dynamic IS equation,

$$\dot{Y}_t^{\text{RA}} = \frac{1}{\gamma}\left(i_t^{\text{RA}} - \pi_t^{w,\text{RA}} + \frac{\dot{A}_t}{A_t} - \rho_t \right) Y_t^{\text{RA}}$$

and the New Keynesian Phillips curve

$$\dot{\pi}_t^{w,\text{RA}} = \rho_t \pi_t^{w,\text{RA}} + \frac{\epsilon_t}{\delta}\left[\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) Y_t^{\text{RA}} u'(Y_t^{\text{RA}}) - v'\left(\frac{Y_t^{\text{RA}}}{A_t}\right)\frac{Y_t^{\text{RA}}}{A_t} \right].$$

In the RANK limit, the "micro block" of our HANK model collapses to a single equation, the aggregate Euler or dynamic IS equation. The macro block takes the same form in both models. However, unions' perceived marginal value of higher wages is proportional to $u'(Y_t^{\text{RA}})$ in RANK, whereas it is governed by the earnings-weighted average of marginal utilities $\iint zu'(c_t(a,z))g_t(a,z)\,da\,dz$ in HANK.

[15] As $\delta \to 0$, equilibrium requires that $0 = \frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)w_t\Lambda_t - v'(N_t)$. That is, the augmented labor wedge is 0 in the flexprice allocation, and we obtain natural output from this equation.

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 the gains from commitment in Section 4.

Ramsey problem with finite commitment horizon. In the discrete-time analysis of Clarida et al. (1999), a planner under discretion chooses policy in each period t taking as given future policy, starting from period $t+1$. Without commitment, the planner cannot credibly influence agents' expectations about the future and so takes these as given as well. Hence, "policy under discretion" is given by the solution to the sequence of these one-period planning problems—formally, a Markov perfect equilibrium of the game played by planners that exercise control for only one period.

We try to stay as close as possible to this notion of policy under discretion. We consider a planner with control over policy in the present who takes future policy—under the control of a future planner—as well as agents' expectations about the future as given. In continuous time, however, we can no longer simply associate the present and future with periods t and $t+1$. To that end, we introduce a *finite-horizon Ramsey problem*. 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*, the present planner is replaced with 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, and we assume that transitions occur at the transition rate ψ.

Following a primal approach, we now formally define the problem of a Ramsey planner who has commitment over the time horizon $[0, \tau_1)$ but takes as given policy enacted by future planners. In general, a Ramsey planner chooses policy in order to maximize a particular objective subject to a set of conditions that define equilibria. Here, we assume that the planner seeks to maximize a utilitarian social welfare function. Lemma 13 in the Appendix formally describes the set of implementability conditions that constrain the Ramsey problem in primal form. The finite-horizon Ramsey problem we now introduce is useful because it allows us to study policy with and without commitment in a unified framework, as we show in remainder of Sections 3 and 4.[16]

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

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

as well as multipliers

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

[16] Formally, our derivation merges insights from Marcet and Marimon (2019) with the continuous-time approach of Harris and Laibson (2013). See Appendix E for details.

to maximize social welfare subject to implementability conditions (8, 15 – 18), taking as given the initial cross-sectional distribution $g_0(a,z)$ as well as future policy.[17] *A Ramsey problem with commitment horizon $[0, \tau_1)$ therefore solves*

$$\mathcal{W}_0(g_0) = \min_{\boldsymbol{M}} \max_{\boldsymbol{X}} \mathbb{E}_0 \left[L(0, \tau_1, g_0) + e^{-\int_0^{\tau_1} \rho_s ds} \mathcal{W}_{\tau_1}(g_{\tau_1}) \right] \tag{20}$$

where the expectation $\mathbb{E}_0$ is over the transition time τ_1, $\mathbb{E}_0 \left[e^{-\int_0^{\tau_1} \rho_s ds} \mathcal{W}_{\tau_1}(g_{\tau_1}) \right]$ 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 $g_{t_1}(a,z)$:

$$\begin{aligned} L(t_1, t_2, g_{t_1}) = \int_{t_1}^{t_2} e^{-\int_0^t \rho_s ds} \Bigg\{ \iint \Bigg\{ & U_t(a,z)\, g_t(a,z) \\ & + \phi_t(a,z) \Big[-\rho_t V_t(a,z) + U_t(a,z) + \partial_t V_t(a,z) + \mathcal{A}_t V_t(a,z) \Big] \\ & + \chi_t(a,z) \Big[u'(c_t(a,z)) - \partial_a V_t(a,z) \Big] \\ & + \lambda_t(a,z) \Big[-\partial_t g_t(a,z) + \mathcal{A}_t^* g_t(a,z) \Big] \Bigg\} \, da\, dz \\ & + \mu_t \Big[\iint (c_t(a,z) - A_t z N_t) g_t(a,z) \, da\, dz \Big] \\ & + \theta_t \Big[-\partial_t \pi_t^w + \rho_t \pi_t^w + \frac{\epsilon_t}{\delta} \iint \tau_t(a,z) g_t(a,z) \, da\, dz \Big] \Bigg\} dt, \end{aligned} \tag{21}$$

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.

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.

Formally, for a given ψ, we study the Markov perfect equilibrium of the game played by a

[17] 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.

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 (8, 15 – 18) 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, $\phi_t(a,z)$, $\chi_t(a,z)$, $\lambda_t(a,z)$, μ_t, and θ_t that solve (20). Policy under discretion corresponds to the limit of this equilibrium as $\psi \to \infty$ and planners have vanishingly small commitment horizons.

The implementability constraints encoded in the Lagrangian (21) incorporate all of the conditions that characterize competitive equilibrium. The first three sets of constraints map to equations (15) through (17) and apply to every pair of individual states (a,z). The last two constraints are of aggregate nature, representing the aggregate resource constraint (18) and the New Keynesian wage Phillips curve (8). Note that this is a minimal characterization of the Ramsey problem in primal form, in the sense that there are no redundant constraints that can be easily deduced from other constraints.[18]

Two of these implementability conditions are forward-looking constraints: the individual Bellman equation and the Phillips curve, respectively associated with the multipliers $\phi_t(a,z)$ and θ_t. A planner under discretion takes future policy—and consequently agents' expectations about future policy, allocations, and prices—as given. In Clarida et al. (1999)'s discrete-time analysis of the representative-agent New Keynesian model, lack of commitment implies the planner takes as given expectations over next period's inflation. In the continuous-time problem (20), the planner similarly takes as given expectations about inflation and value assignments in the future, i.e., 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.

Policy under discretion: optimality conditions. 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 (20) as $\psi \to \infty$.

Proposition 1. (**Policy under Discretion: Optimality Conditions**) *The necessary first-order conditions*

[18] We could have defined an alternative Ramsey problem that includes households' consumption Euler equations instead of their HJB equations and consumption optimality conditions. As will become clear below, our formulation provides valuable analytical insights, by virtue of having two sets of multipliers associated with households' lifetime value and consumption-savings decisions, $\phi_t(a,z)$ and $\chi_t(a,z)$. Our formulation is also computationally more tractable.

that characterize optimal monetary policy under discretion are given by

$$\rho_t \lambda_t(a,z) = U_t(a,z) + \mu_t(c_t(a,z) - A_t z N_t) + \partial_t \lambda_t(a,z) + \mathcal{A}_t \lambda_t(a,z) \tag{22}$$

$$u'(c_t(a,z)) = \partial_a \lambda_t(a,z) - \mu_t + \tilde{\chi}_t(a,z) \tag{23}$$

$$0 = \iint z \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{z}\, \zeta_t^{HTM} g_t(\underline{a},\underline{z})\, da\, dz - \mu_t - \frac{v'(N_t)}{A_t} \tag{24}$$

$$0 = \iint a \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{a}\, \zeta_t^{HTM} g_t(\underline{a},\underline{z})\, da\, dz \tag{25}$$

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 (23) through (25) correspond to the first-order conditions for individual consumption $c_t(a,z)$, the cross-sectional distribution, $g_t(a,z)$, aggregate activity, N_t, and the nominal interest rate, i_t, respectively. By judiciously combining these conditions, we will be able 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. The multiplier $\chi_t(a,z)$ corresponds to the social shadow value of relaxing the individual consumption-savings decision of households in state (a,z) in the direction of consuming more. 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 fully manage consumption-savings decisions.[19] The multiplier $\lambda_t(a,z)$ corresponds to the social shadow value of increasing the mass of households in state (a,z). We show below that it represents the *social lifetime value* of a household in state (a,z). Finally, 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. Under discretion, this occurs when the economy is in a recession (boom) and the aggregate labor wedge is positive (negative) or, equivalently, the output gap is negative (positive).

Next, we interpret the four optimality conditions (22) – (25). First, equation (22) defines $\lambda_t(a,z)$ as the social lifetime value of a household. It 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). The difference between private lifetime value (15) and social lifetime value (22) under discretion is given by the term $\mu_t(c_t(a,z) - w_t z N_t)$, which captures the contribution of a household in state (a,z) to aggregate excess demand. Intuitively, households for whom $c_t(a,z) > A_t z N_t$

[19] If $\chi_t(a,z)$ were 0 for all (a,z) and t, the planner would agree with households on their private consumption-savings decisions. A non-zero $\chi_t(a,z)$ then represents the shadow penalty that ensures the planner respects individual behavior. Nuño and Thomas (2020) study a small open economy environment where $\chi_t(a,z) = 0$.

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, $w_t z N_t$. When $\mu_t > 0$, i.e., increasing aggregate excess demand is social valuable (the economy is in a recession), the planner will attach a higher social value to these households. The opposite occurs when $\mu_t < 0$.

Equation (22) allows us to characterize the social marginal value of wealth—a key input for the remaining optimality conditions—as

$$\partial_a \lambda_t(a,z) = \partial_a V_t(a,z) + \mathcal{M}_t(a,z)\mu_t, \tag{26}$$

where $\mathcal{M}_t(a,z)$ denotes an operator that acts on the path of multipliers μ_t. In economic terms, $\mathcal{M}_t(a,z)$ corresponds to the present discounted value of marginal propensities to consume, which we make formal in Lemma 36 in Appendix E.[20] The difference between the private and the social marginal 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. Intuitively, a marginal increase in wealth translates into higher aggregate demand at time t and in the future, depending on the household's propensities to consume and save out of wealth. Such spending is socially beneficial when $\mu_t > 0$ or costly when $\mu_t < 0$—an effect that only the planner internalizes. Note that a planner under discretion accounts for the social impact of future consumption via the path of future multipliers μ_t, despite taking future policy and expectations as given.

Second, equation (23) has the interpretation of a social consumption-savings optimality condition for households in state (a,z). Similarly to households, the planner trades off the direct benefit of increasing consumption, $u'(c(a,z))$, 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 beneficial when $\mu_t > 0$ (recession) or costly when $\mu_t < 0$ (boom). Finally, since the planner must respect private consumption-savings decisions, she also faces a social shadow cost from adjusting consumption, $\tilde{\chi}_t(a,z)$. Consistently with the interpretation of $\tilde{\chi}_t(a,z)$ above, when the planner perceives that a household is consuming too little/saving too much, $\tilde{\chi}_t(a,z) > 0$ acts as an additional shadow social cost of consumption, so that the individual consumption-savings optimality condition is satisfied. Instead, when a planner perceives that a households is consuming too much/saving too little, $\tilde{\chi}_t(a,z) < 0$ acts as an additional shadow social benefit of consumption.

Before moving to interpreting equations (24) and (25) it is valuable to leverage our understand-

[20] We show there that $\mathcal{M}_t(a,z)$ is positive and bounded between 0 and 1 under some mild regularity conditions. In particular, the operator admits the representation

$$\mathcal{M}_t(a,z) = (\rho - r_t + \partial_a c_t(a,z) - \partial_t - \mathcal{A}_t)^{-1} \partial_a c_t(a,z).$$

The term $\rho - r_t + \partial_a c_t(a,z) = \rho - \partial_a s_t(a,z)$ captures time discounting net of the interest rate on the assets not consumed. The terms ∂_t and $\mathcal{A}_t$ account for changes in aggregate conditions over time, ∂_t, and for the expected transition of the household across states, $\mathcal{A}_t$. Finally, $\partial_a c_t(a,z)$ is simply the instantaneous marginal propensity to consume (MPC).

ing of $\chi_t(a,z)$ and $\partial_a \lambda_t(a,z)$ to determine how the planner values the consumption of a household at the borrowing constraint. In particular, it useful to characterize $\partial_a \lambda_t(\underline{a},\underline{z}) g_t(\underline{a},\underline{z}) + \xi_t^{\text{HTM}}$, given by

$$\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t^{\text{HTM}} = u'(c_t(\underline{a},\underline{z})) + \mu_t \tag{27}$$

which captures the social value of increasing the consumption of constrained households. The planner must respect the borrowing constraint households face. She internalizes, therefore, that a marginal change in the income of a household at the borrowing constrained leads, one-for-one, to a change in consumption. The planner's perceived social marginal value of wealth for the constrained household is therefore equal to the planner's valuation of a marginal increase in consumption, which comprises the sum of the direct utility benefit, $u'(c_t(\underline{a},\underline{z}))$, and the impact on aggregate excess demand.

Third, equation (24) represents the planner's valuation of an increase in hours worked by all households.[21] Such an increase in hours comprises three effects. 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 hand-to-mouth households, $\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t(\underline{a},\underline{z})$. Second, aggregate supply increases by $\iint zA_t g_t(a,z)\, da\, dz = A_t$—valued at the shadow value of aggregate excess demand, μ_t—which is socially beneficial when the economy is overheated. Finally, households experience a 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 effects.

Finally, equation (25) 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 according to $\partial_a \lambda_t(\underline{a},\underline{z}) + \xi_t^{HTM}$ for constrained households. We expect the planner to find it desirable to redistribute resources from households who save, with $a > 0$, and have a low social marginal value of wealth, to households who borrow, with $a < 0$, and have a high social marginal value of wealth. These terms capture the distributive pecuniary effect of a change in interest rates, which is central to the determination of optimal monetary policy as we show next.

Targeting rule and inflationary bias. Combining these optimality conditions allows us to characterize a *targeting rule* for optimal monetary policy under discretion in Proposition 2 below. This targeting rule clarifies the tradeoffs optimal monetary policy under discretion balances in our HANK model. It also 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.

It is instructive to derive this targeting explicitly in the main text. To that end, we introduce

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

two useful policy perturbations that are purposefully designed so that they have a neutral impact on aggregate excess demand. The first perturbation combines equations (23) and (24), yielding an *aggregate activity condition*. This perturbation entails making households work an extra hour while forcing them to consume the proceeds of any additional income. At an optimum, the marginal value of this perturbation for a planner must satisfy

$$\underbrace{\iint \left(zu'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) g_t(a,z)\, da\, dz}_{\text{Aggregate Labor Wedge}} - \iint z\tilde{\chi}_t(a,z) g_t(a,z)\, da\, dz = .$$

As one would expect, the aggregate labor wedge captures the social marginal benefit of increasing aggregate activity. And if the planner had the ability to control households consumption-savings decisions (i.e., if $\chi_t(a,z) = 0$), the aggregate activity condition shows that a planner would set the labor wedge to zero. However, the planner must account for the fact that increasing consumption impacts households' savings decisions, which counterbalacnes the desire to set the aggregate labor wedge to zero.[22]

The second perturbation combines equations (23) and (25) and yields an *interest rate condition*. This perturbation entails a unit increase in interest rates while making households directly consume the resulting pecuniary gains. At an optimum, the marginal value of this perturbation for a planner must satisfy

$$\underbrace{\iint au'(c_t(a,z)) g_t(a,z)\, da\, dz}_{\text{Distributive Pecuniary Effect}} - \iint a\tilde{\chi}_t(a,z) g_t(a,z)\, da\, dz = 0$$

In this case, the social marginal benefit of increasing interest rates is captured by its distributive pecuniary effect, which is negative.

[23]

The planner understands that a change in rates simply redistributes resources across savers and borrowers, since distributive pecuniary effects are always zero-sum in aggregate in dollar terms. However, since borrowers typically have a higher marginal utility of consumption than savers, a utilitarian planner perceives that an increase in rates decreases social welfare through this channel. This desire to redistribute towards high marginal utility households by reducing interest rates—a motive that is absent in representative-agent economies—is a central determinant of optimal

[22] The aggregate activity condition also connects μ_t directly to the aggregate labor wedge when policy is set with discretion. When substituting in for $\tilde{\chi}_t(a,z)$ from equation (23), we obtain a condition that defines μ_t as a weighted sum of future labor wedges:

$$\mu_t = \frac{\iint \left(zu'\left(c_t\left(a,z\right)\right) - \frac{v'(N_t)}{A_t} \right) g_t\left(a,z\right)\, da\, dz}{\iint z\left(1 - \mathcal{M}_t\left(a,z\right)\right) g_t\left(a,z\right)\, da\, dz}$$

[23] We use the terminology distributive pecuniary externalities 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 implies that distributive pecuniary effects are zero, so $\iint a\partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz = 0$.

monetary policy in our environment. As in the case of the aggregate activity perturbation, the planner must account for the fact that change interest rates impacts households' savings decisions.

While both policy perturbations are neutral in terms of aggregate excess demand, they are not neutral intertemporally in terms of their impact on households' savings decisions. However, we can scale and combine both perturbations to neutralize the intertemporal effect, obtaining the targeting rule of Proposition 2. This targeting rule shows that, under discretion, a utilitarian planner in a heterogeneous-agent environment trades off aggregate stabilization against redistribution.

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

$$\underbrace{\iint \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) g_t(a,z)\, da\, dz}_{\textit{Aggregate Labor Wedge}} = \Omega_t^{\mathrm{D}} \underbrace{\iint a u'(c_t(a,z)) g_t(a,z)\, da\, dz}_{\textit{Distributive Pecuniary Effect}} \tag{28}$$

where

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

Equation (28) shows that the welfare impact of a perturbation that jointly increases hours worked by all households and reduces the interest rate by Ω_t^D basis points—forcing households to consume the resulting proceeds—must be zero at an optimum.[24] Equation (28) implies that, at an optimum, the planner sets policy trading off aggregate stabilization and redistribution motives.

The LHS of equation (28) is the aggregate labor wedge and represents the aggregate stabilization motive of the planner. The RHS represents the redistribution motive of the planner. 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 $\mathbb{C}ov_{g_t(a,z)}(a, u'(c_t(a,z)))$ denotes the cross-sectional covariance between wealth and marginal utility at time t. The redistribution term on the RHS of equation (28) is consequently negative.

Proposition 2 therefore implies that, in the presence of a redistribution motive, a utilitarian planner under discretion targets a negative aggregate labor wedge, implying an overheated economy. Proposition 2 thus offers a novel perspective on optimal monetary policy under discretion in heterogeneous-agent environments.

In the RANK limit of our economy, the planner's motive to redistribute via distributive pecuniary interest rate effects naturally vanishes. Formally, the RHS of equation (28) 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.

[24] And since the net present value MPC, $\mathcal{M}_t$, is bounded between 0 and 1, we have $\Omega_t^{\mathrm{D}} > 0$.

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

$$Y_t = \tilde{Y}_t \times \underbrace{\left(\frac{\epsilon_t}{\epsilon_t - 1}\frac{1}{1+\tau^L}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Desired Markups: } \geq 1} \times \underbrace{\left(1 - \Omega_t^{\mathrm{D}} \frac{\iint au'(c_t(a,z))g_t(a,z)\,da\,dz}{\iint zu'(c_t(a,z))g_t(a,z)\,da\,dz}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Benefit from Redistribution: } > 1} \tag{29}$$

where $\tilde{Y}_t$ denotes natural output as defined in equation (19).

While this output gap targeting rule connects more directly to the canonical results on optimal monetary policy in RANK, equations (28) and (29) have the same content.

Equation (29) 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. As in RANK, this distortion motivates the planner to raise output above potential, i.e., overheat the economy to raise employment. This desired markup wedge is positive whenever the employment subsidy τ^L is not sufficiently large. In the RANK limit of our economy,where $\Omega_t^{\mathrm{D}} \to 0$, the desired markup wedge is the only motive for the planner to target a non-zero output gap under discretion.

In HANK, a second redistribution wedge emerges. Since marginal utility of consumption falls with wealth, we have $\mathbb{C}ov_{g_t(a,z)}(a, u'(c_t(a,z))) = \iint au'(c_t(a,z))g_t(a,z)\,da\,dz < 0$. The redistribution wedge is therefore strictly positive, encouraging the utilitarian planner under discretion to overheat the economy even further.

An important conclusion of the canoncial 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. In that case, with $\Omega_t^D \to 0$ in the RANK limit and $\frac{\epsilon_t}{\epsilon_t-1}\frac{1}{1+\tau^L} = 1$, the targeting rule (29) simply collapses to $Y_t = \tilde{Y}_t$: the target of monetary policy under discretion is then to close the output gap, which in this case also closes the inflation gap. 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.

Without commitment, the targeting rule (29) suggests that the planner always faces an incentive to raise output above natural output or, by equation (28) engineer a negative labor wedge, i.e., a boom. We next investigate the long-term consequences of policy under discretion. Consider a stationary point of the Markov perfect equilibrium for policy under discretion. The following Proposition shows that such a steady state features substantially larger inflationary bias than in the RANK limit.

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

$$\pi^w_{ss} = \frac{\epsilon}{\delta} A_{ss} N_{ss} \Bigg[\underbrace{\left(1 - \frac{\epsilon - 1}{\epsilon}(1+\tau^L)\right)\Lambda_{ss}}_{\textit{Monopolistic Competition: } \geq 0} - \Omega^D_{ss} \underbrace{\mathbb{C}ov_{g_{ss}(a,z)}\Big(a, u'(c_{ss}(a,z))\Big)}_{\textit{Redistribution: } < 0} \Bigg]. \tag{30}$$

Inflationary bias now emerges from two sources: inefficiently low steady state employment due to monopolistic competition and redistribution. Quantitatively, the contribution of the redistribution motive is over 4 times larger than that of monopolistic competition in our calibration exercise (also see Figure 1 in Section 4.3). So the redistribution incentive of a utilitarian planner under discretion contributes more than 4 times as much to steady state inflationary bias as the classical effect due to Barro and Gordon (1983).

Practical implications of optimal monetary policy under discretion. There are three main takeways from our analysis so far. First, when monetary policy is set by a utilitarian planner under discretion, the economy runs overheated. The utilitarian planner values redistribution towards high-marginal utility households and trades off this redistribution motive against aggregate stabilization.

Second, the stationary equilibrium with policy under discretion features inflationary bias in the sense of Barro and Gordon (1983). In our setting, inflationary bias results not only from the standard motive owing to monopolistic competition but is further exacerbated by the redistribution motive: When setting policy under discretion, the desire to redistribute towards high marginal utility households manifests in form of inflation. Quantitatively, this redistribution motive is substantially larger than the standard source of inflationary bias.

Third, an appropriate employment subsidy—set so that $\frac{\epsilon-1}{\epsilon}(1+\tau^L) = 1$—can eliminate the steady state markup distortion that stems from monopolistic competition. As in RANK, the employment subsidy can fully address this source of inflationary bias. In HANK, however, an employment subsidy alone cannot fully undo inflationary bias in steady state due to the redistribution motive. Gains from commitment remain large.

4 A Timeless Ramsey Approach

We have shown in Section 3 that the redistribution motive exacerbates inflationary bias when a utilitarian planner sets policy under discretion. In this section, we develop a *timeless Ramsey approach* to characterize optimal monetary policy under commitment in heterogeneous-agent economies.

Our approach proceeds in three steps. First, in Section 4.1, we define a standard Ramsey problem, which in turn allows us to define and characterize Ramsey plans and stationary Ramsey plans. In particular, we show in Section 4.2 that the optimal stationary equilibrium under commitment features 0 inflation, addressing the inflationary bias of policy under discretion in the long run.

Second, we show in Section 4.3 that, while the standard Ramsey problem addresses inflationary bias in the long run, it suffers from multiple dimensions of time inconsistency and consequently still features inflationary bias in the short run. In order to find time-consistent planning solutions, 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 is time-consistent. That is, the timeless Ramsey planner no longer has an incentive to deviate from the stationary Ramsey plan in the absence of shocks. The timeless Ramsey problem resolves inflationary bias in both the short run and the long run. In Section 4.4, we associate the timeless penalties required for time-consistent monetary policy in our model with an inflation target in the spirit of Barro and Gordon (1983) and a novel distributional target.

Third, in Section 4.5, we characterize optimal stabilization policy under the timeless Ramsey problem.

Each of the three steps of the timeless Ramsey approach allows us to isolate one important dimension of optimal monetary policy design. Characterizing the stationary Ramsey plan allows us to solve for optimal long-run policy, with which the planner addresses distortions and inefficiencies she perceives in a stationary equilibrium. 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 time consistency problems. And finally, by characterizing optimal stabilization policy with the appropriate timeless penalty around the stationary Ramsey plan, we isolate the planner's pure stabilization motive and no longer confound it with considerations of long-run distortions and time inconsistency.

4.1 Standard Ramsey Problem and Ramsey Plan

The standard Ramsey problem corresponds to the limit of problem (20) as $\psi \to 0$, i.e., as the commitment horizon becomes infinite.

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

a) *A standard primal Ramsey problem solves*

$$\min_{\{\phi_t(a,z), \chi_t(a,z), \lambda_t(a,z), \mu_t, \theta_t\}} \max_{\{c_t(a,z), V_t(a,z), g_t(a,z), N_t, \pi_t^w, i_t\}} L^{\text{SP}}(g_0) \tag{31}$$

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

$$L^{\text{SP}}(g_0) = \lim_{T\to\infty} L(0, T, g_0). \tag{32}$$

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 competitive equilibrium 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* (31)*.*

We define the standard Ramsey problem (31) as the limit of problem (20) as the commitment horizon becomes infinite. For convenience, we state the full problem in Appendix A and emphasize here that problem (31) corresponds exactly to the conventional Ramsey problem in primal form as defined, e.g., in Ljungqvist and Sargent (2018).

In Proposition 4 we summarize the optimality conditions that characterize the standard Ramsey plan. Our derivation relies on a variational approach, formally developed in Appendix A.

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

$$\partial_t \phi_t(a,z) = -\mathcal{A}_t^* \phi_t(a,z) + \partial_a \chi_t(a,z) \tag{33}$$

$$\rho_t \lambda_t(a,z) = U_t(a,z) + \mu_t(c_t(a,z) - A_t z N_t) + \theta_t \frac{\epsilon_t}{\delta} \tau_t(a,z) + \partial_t \lambda_t(a,z) + \mathcal{A}_t \lambda_t(a,z) \tag{34}$$

$$u'(c_t(a,z)) = \partial_a \lambda_t(a,z) - \mu_t - \theta_t \frac{\epsilon_t}{\delta} \frac{d\tau_t(a,z)}{dc_t(a,z)} + \tilde{\chi}_t(a,z) \tag{35}$$

$$\begin{aligned} 0 = & \iint z \partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz + \underline{z} \xi_t^{HTM} g_t(\underline{a}, \underline{z})\, da\, dz - \mu_t - \frac{v'(N_t)}{A_t} \\ & + \iint \phi_t(a,z) \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) da\, dz + \theta_t \frac{\epsilon_t}{\delta} \iint \frac{1}{A_t} \frac{d\tau_t(a,z)}{dN_t} g_t(a,z)\, da\, dz \end{aligned} \tag{36}$$

$$\begin{aligned} \dot{\theta}_t = & \delta \pi_t^w \left(1 + \iint \phi_t(a,z)\, da\, dz \right) \\ & + \iint \left(a \partial_a \lambda_t(a,z) g_t(a,z) + a \phi_t(a,z) \partial_a V_t(a,z) \right) da\, dz + \underline{a} \xi_t^{HTM} g_t(\underline{a}, \underline{z}) \end{aligned} \tag{37}$$

$$0 = \iint \left(a \partial_a \lambda_t(a,z) g_t(a,z) + a \phi_t(a,z) \partial_a V_t(a,z) \right) da\, dz + \underline{a} \xi_t^{HTM} g_t(\underline{a}, \underline{z}) \tag{38}$$

where[25]

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

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

$$0 = \theta_0 \tag{39}$$

$$0 = \phi_0(a,z). \tag{40}$$

The optimality conditions (33) – (34) hold everywhere in the interior of the idiosyncratic state space. For a formal treatment of boundary conditions, see Appendices A.3 through A.6.

Equations (33), (35), (34), (36), and (37) respectively correspond to the optimality conditions for the value function, consumption, the cross-sectional distribution of bond holdings and idiosyncratic labor productivities, aggregate labor, and wage inflation. Finally, equation (38) corresponds to the optimality condition for the nominal interest rate, i.e., monetary policy. We present the proof of Proposition 4 first in continuous time in Appendix A.2 for the interior of the idiosyncratic state space. We then provide a formal treatment of boundary conditions in Appendices A.3 through A.6.

The optimality conditions for a standard Ramsey planner with commitment can be seen as an augmented version of the optimality conditions for a discretionary planner, described in Section 3. Building on our analysis of policy under discretion, we present new insights that emerge from the six optimality conditions of the Ramsey problem.

First, the multipliers on the forward-looking constraints, $\phi_t(a,z)$ for households' lifetime utility and θ_t for the Phillips curve, have the dual interpretation of promises that a planner must satisfy or penalties that a planner faces when choosing households' utility or inflation, respectively. Throughout the paper, we refer to θ_t as an inflation penalty or target, and to $\phi_t(a,z)$ as a distributional penalty or target. In fact, by setting $\phi_t(a,z) = 0$ and $\theta = 0$, and dropping the two equations that define the laws of motion by both multipliers, we recover the optimality conditions for policy under discretion.

Second, the multipliers $\chi_t(a,z)$, $\lambda_t(a,z)$, and μ_t have the same interpretation as in the problem with discretion. Namely, $\chi_t(a,z)$ corresponds to the social shadow value of relaxing the individual consumption-savings decision, with $\chi_t(a,z) > 0$ when households over-borrow and $\chi_t(a,z) < 0$ when households over-save from the planner's perspective; $\lambda_t(a,z)$ is the social lifetime value of a household in state (a,z); and μ_t corresponds to the social shadow value of increasing aggregate excess demand.

[25] The interpretation of ζ_t^{HTM} is the social marginal value of giving a dollar of (unearned) income to every hand-to-mouth household. For the perturbations where ζ_t^{HTM} shows up, 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.

Third, equations (34), (35), (36), and (38) are the counterparts of equations (22) through (25). Here, they are augmented by the penalties necessary for a planner to sustain previously made promises under commitment. Equation (35) is the counterpart of equation (23), and corresponds to the social consumption-savings optimality condition. With commitment, the planner perceives an additional cost of increasing consumption for households in state (a, z) when increasing consumption reduces the labor wedge, i.e., $\frac{d\tau_t(a,z)}{dc_t(a,z)} < 0$, since this generates inflationary pressure, which is costly for the planner according to the penalty θ_t.

Equation (34) is the counterpart of equation (22). With commitment, the difference between private lifetime value (15) and social lifetime value (34) also includes an additional flow cost associated with the individual labor wedge. Intuitively, households for whom $\tau_t(a, z) > (<)\ 0$ put positive (negative) pressure on inflation, since the contribute positively (negatively) to the aggregate labor wedge. This inflationary pressure is costly for the planner according to the penalty θ_t.

Equation (36) is the counterpart (24) and represents the planner's valuation of an increase in hours worked by all households. With commitment, the planner perceives an additional marginal benefit of increasing hours when $\phi_t(a, z) > 0$ and that household's individual labor wedge is is positive, $zu'(c_t(a, z)) - \frac{v'(N_t)}{A_t} > 0$. The individual labor wedge captures the increase in lifetime utility (decrease in lifetime utility if the labor wedge is negative) that a paricular household experiences when working an extra hour. When $\phi_t(a, z) > 0$ the planner faces a negative penalty (a bounty) for increasing the utility of households in this state, which makes increasing hours more desirable. Furthermore, the planner understands that increasing hours may directly cause inflationary pressure through a change in the aggregate labor wedge. When $\theta_t > 0$, this makes it less desirable to increase aggregate activity. This equation is the foundation of our discussion of optimal stabilization policy in Section 4.5.

Equation (38) is the counterpart of (25). It represents the planner's valuation of an increase in the nominal interest rate. With commitment, the planner values the redistribution of dollars towards households in proportion to $\partial_a \lambda_t(a, z) + \phi_t(a, z)$. That is, besides valuing gains/losses at the social marginal value of wealth, the increase in rates changes the lifetime utility of households by $a\partial_a V_t(a, z)$, and this change in utility faces a negative penalty (bounty) given by $\phi_t(a, z)$.

Fourth, equation (33), which is central to this paper, has the form of a Kolmogorov forward equation augmented to account for births and deaths. This equation tightly connects the multipliers on households' lifetime utility $\phi_t(a, z)$, which encode how all lifetime policies determine households' lifetime utility with the multipliers on households' optimal consumption-savings decisions, $\chi_t(a, z)$, which capture whether the planner perceives that a household is currently over- or under-saving. Intuitively, equation (33) is the key promise-keeping relation that a Ramsey planner with commitment must satisfy. In particular, the evolution of the distribution of distributional penalties must be consistent with the evolution of households across idiosyncratic states, via $\mathcal{A}^*$, but also account for the birth of promises, captured by the term $\partial_t \chi_t(a, z)$.

Fifth, note that the optimality condition for inflation (37) simplifies to

$$\dot{\theta}_t = \delta \pi_t^w \tag{41}$$

whenever the average of the distributional penalties adds up to zero, that is,

$$\iint \phi_t(a,z)\, da\, dz = 0, \ \forall t. \tag{42}$$

While we have formally proven that this condition is satisfied at a stationary Ramsey plan (defined below), we strongly conjecture that it holds everywhere, as we have verified numerically. Moreover, whenever the distributional penalties cancel on average, equation (41) will be identical to the analogous equation in RANK, as we show in Appendix D. In turn, equation (42) has the interpretation that a planner does not over- or under-promise on aggregate in terms of lifetime utilities.

4.2 Optimal Long-Run Policy

We are now ready to characterize the gains from commitment in our environment and ask to what extent Ramsey policy under commitment addresses the inflationary bias associated with discretion. To that end, we formally define a stationary Ramsey plan, towards which a Ramsey plan may converge when all shocks converge as $t \to \infty$. Finding a stationary Ramsey plan is critical not only to characterize optimal long-run policy but also to be able to leverage perturbation methods to compute optimal stabilization policy.

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.*

In Lemma 21 in the Appendix, we describe the conditions that explicitly characterize a stationary Ramsey plan. They in turn derive from our characterization of the Ramsey plan in Proposition (4).[26]

What are the implications of household heterogeneity for optimal long-run inflation? When policy is set with discretion, the planner's redistribution motive substantially exacerbates inflation-

[26] From the stationary counterpart of equation (33), it immediately follows that there is a key necessary condition for the existence of a stationary Ramsey plan, given by $\iint \partial_a \chi_{ss}(a,z)\, da\, dz = 0$. We highlight this condition because it has an important economic interpretation in the context of equation (33). It implies that the "births" and "deaths" of distributional promises must average out to zero in a stationary Ramsey plan. In the Appendix, we show that this condition is satisfied in our baseline HANK model. This result has the interpretation that a planner does not want to over- or under-promise in the aggregate in terms of lifetime utilities.

ary bias. The following Proposition shows that optimal long-run inflation under commitment is zero even in the presence of household heterogeneity, i.e., the stationary Ramsey plan features 0% inflation.

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. Inflation and the nominal interest rate affect households' financial income, which itself is proportional to the real interest rate $r_{\text{ss}} = i_{\text{ss}} - \pi_{\text{ss}}$, symmetrically. 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. Mathematically, any stationary Ramsey plan must feature $\dot{\theta}_{\text{ss}} = 0$ since θ_{ss} is constant. Equation (41) then directly implies that optimal long-run inflation is zero, $\pi_{\text{ss}} = \pi_{\text{ss}}^{w} = 0$.

It is important to emphasize that the planner can only maintain zero inflation in the long run under commitment. As we discuss in Section 3, the planner always faces an incentive to generate surprise inflation. Even with the appropriate employment subsidy to address the monopolistic markup distortion, a utilitarian planner has the incentive to use surprise inflation for redistribution. Under discretion, such surprise inflation is self-defeating as it simply results in inflationary bias. It is only with commitment that the planner can promise not to generate surprise inflation ex-post.

We conclude our discussion of optimal long-run policy with several remarks. First, the Ramsey optimality condition for inflation in RANK, equation (73), is exactly parallel to that in HANK, equation (37). In the RANK limit, the optimality condition for the interest rate collapses to $\phi_t^{\text{RA}} = 0$, which automatically removes any desire to adjust policy to change the (degenerate) distribution of consumption promises. The Ramsey planner in principle also has a distributional "target" as in HANK but this target becomes degenerate in the representative-agent limit. Consequently, the economic determinants of the zero optimal long-run inflation result in both models are the same. Second, given our assumptions, it should not be surprising that zero-inflation is the optimal long-run policy. A rich literature, including Chari and Kehoe (1999), Khan et al. (2003), and Schmitt-Grohé and Uribe (2010), has studied environments in which the optimal long-run policy features non-zero inflation. Third, a corollary of our analysis is that household heterogeneity can only be a source of non-zero optimal long-run inflation if the nominal interest rate and inflation affect individual households asymmetrically. That is, in environments in which inflation and the nominal interest rate have a differential impact on the distribution of individuals in the economy, we should expect an optimal long-run policy that features non-zero inflation.[27]

[27] For example, if households consume different bundles of goods and face different price indices, then inflation will

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

Our definition of the standard Ramsey problem assumes commitment by the planner from time 0 onwards. The resulting Ramsey plan features zero inflation in steady state, thus addressing the problem of inflationary bias in the long run. Importantly, however, standard Ramsey plans will generically be time-inconsistent as we describe next. As a result, standard Ramsey plans still feature inflationary bias in the short run. Motivated by this observation, we introduce and define *timeless penalties* in this subsection and use them to formalize a *timeless Ramsey problem*, which we show is time-consistent and addresses inflationary bias both in the long run and in the short run.

The implementability conditions that constrain the Ramsey planner include two sets of forward-looking conditions: individual Bellman equations and the New Keynesian Phillips curve. Each of these implementability conditions is a source of time inconsistency.

First, consider equation (33), which characterizes the evolution of $\phi_t(a,z)$, the multiplier associated with the forward-looking individual Bellman equation. In our continuous-time setup, time inconsistency materializes as follows: We know from the standard Ramsey problem that optimality requires an initial condition

$$\phi_0(a,z) = 0$$

for all (a,z) because $V_0(a,z)$ is free. But any stationary Ramsey plan will generically feature $\phi_{ss}(a,z) \neq 0$, which follows directly from the stationary version of equation (33), as long as $\partial_a \chi_{ss}(a,z) \neq 0$. 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 generically 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 condition of the standard Ramsey problem that $\phi_0(a,z) = 0$ and precisely formalizes the first form of time inconsistency in our setting.[28]

The second source of time inconsistency is the forward-looking Phillips curve. While the first form of time inconsistency owes to the planner's incentive to redistribute across agents, the second one owe's to the planner's incentive to exploit firms' locked-in prices. Optimality in the standard Ramsey problem requires the initial condition

$$\theta_0 = 0$$

because the planner's choice of inflation at time 0, π_0^w, is free. As in the case of $\phi_{ss}(a,z)$, the stationary Ramsey plan generically features $\theta_{ss} \neq 0$. This is the classical time consistency problem in the standard New Keynesian model studied by the vast literature following Barro and Gordon

have asymmetric effects on households' real purchasing power (Clayton et al., 2018; Cravino et al., 2020).

[28] Note that Acharya et al. (2020) have already identified a version of this form of time inconsistency in a HANK economy. More generally, in incomplete market environments in which pecuniary externalities are present, it is understood that optimal policies are typically time-inconsistent—see, e.g., Bianchi and Mendoza (2018) and Jeanne and Korinek (2020).

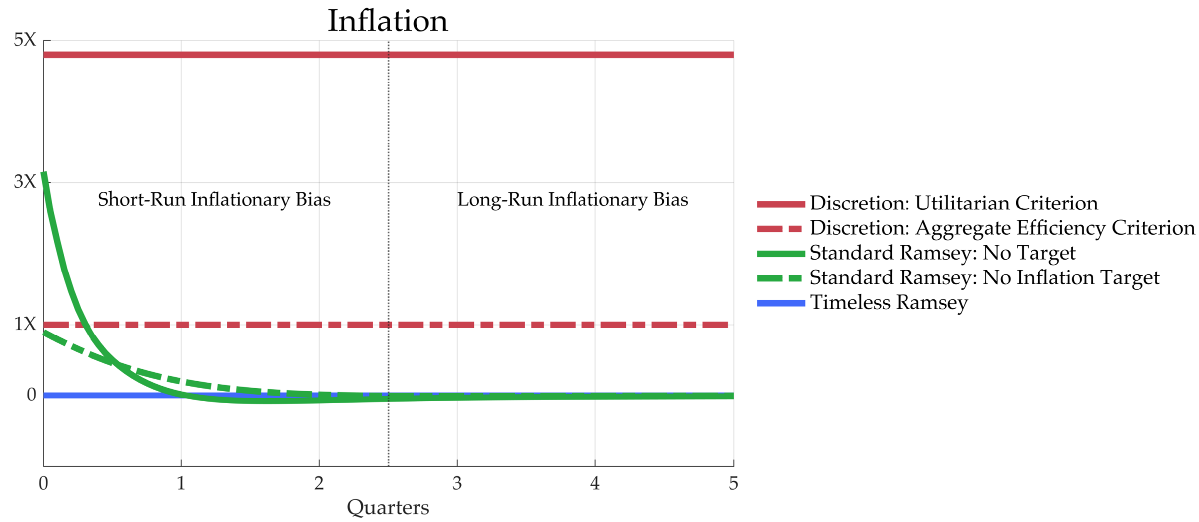


Figure 1. Inflationary Bias in the Short Run and the Long Run

Note. Figure 1 illustrates the inflationary bias associated with different planning problems in the short run and in the long run. We normalize to 1 the inflationary bias associated with policy under discretion when there is no redistribution motive, i.e., the classical inflationary bias associated with monopolistic competition (red, dashed). All units are relative to this baseline. 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 time consistency problems, however, standard Ramsey plans still feature short-run inflationary bias. Only the timeless Ramsey problem (blue), under which monetary policy is time-consistent, fully addresses inflationary bias in both the short run and the long run.

(1983).

These time consistency problems imply that standard Ramsey plans still feature inflationary bias in the short run, as we illustrate in Figure 1. The long-run inflationary bias that plagues the stationary equilibrium under discretion (red lines) is absent from the stationary Ramsey plan under commitment (green lines). However, the standard Ramsey planner has an incentivize to use surprise inflation in the short run because, at time 0, no past promises constrain the planner. Inflationary bias therefore persists in the short run.

Motivated by this observation, we now introduce *timeless penalties*, which are at the heart of our approach. As we show below, a Ramsey planner confronted with these timeless penalties no longer has an incentive to use surprise inflation in the short run. That is, the timeless penalties ensure the planner already internalizes at time 0 the promises she would like to make for the future. Formally, building on Marcet and Marimon (2019), we introduce a penalty or initial promise for each forward-looking implementability condition that the Ramsey planner faces. In Section 4.4, we associate the timeless penalties with an inflation target in the spirit of Barro and Gordon (1983) and a novel distributional target.

Definition 4. (**Timeless Penalty**) *We define the timeless penalty as*

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

where we refer to $\phi(a,z)$ as the initial distributional promise and θ the initial inflation promise.

In principle, the timeless penalty is defined for any pair of initial promises $\phi(a,z)$ and θ. As we show next, when we confront the Ramsey planner with the timeless penalty for a specific set of initial promises, the penalty exactly offsets the two time consistency problems. Intuitively, the timeless penalty ensures that the Ramsey planner at time 0 behaves consistently with the promises that she would like to make for the future. In other words, the timeless penalty encodes precisely those promises that the planner would like to make for the long run under commitment.

In order to characterize optimal time-consistent monetary policy, we now define the timeless Ramsey problem (Woodford, 1999, 2003, 2010) in its primal form.[29, 30] Formally, we build on and extend Marcet and Marimon (2019) to our setting—continuous-time heterogeneous-agent economies—by augmenting the standard Ramsey problem with the timeless penalties defined above. The resulting Ramsey problem is indexed by and defined for any set of initial promises ϕ and θ, which act as state variables. The main result of this section is that, under the appropriate penalty, the timeless Ramsey problem is time-consistent: a timeless Ramsey planner has no incentive to deviate from the stationary Ramsey plan in the absence shocks.

Definition 5. (**Timeless Primal Ramsey Problem**) *A timeless primal Ramsey problem solves*

$$\min_{\{\phi_t(a,z),\,\chi_t(a,z),\lambda_t(a,z),\mu_t,\theta_t\}} \; \max_{\{c_t(a,z),V_t(a,z),g_t(a,z),N_t,\pi_t^w,i_t\}} L^{\mathrm{TP}}(g_0,\phi,\theta),$$

where $L^{\mathrm{TP}}(g_0,\phi,\theta)$ denotes the timeless primal Lagrangian, given an initial distribution g_0 as well as two new state variables: a distributional promise ϕ and an inflation promise θ. The Lagrangian is defined as

$$L^{\mathrm{TP}}(g_0,\phi,\theta) = L^{\mathrm{SP}}(g_0) + \mathcal{T}(\phi,\theta) \tag{43}$$

where $L^{\mathrm{SP}}(g_0)$ is the standard primal Lagrangian defined in equation (32).

While we define the Lagrangian $L^{\mathrm{TP}}(g_0,\phi,\theta)$ for an arbitrary set of initial promises ϕ and θ, it will become clear in the following that the timeless primal Ramsey problem is time-consistent only if we set $\phi(a,z) = \phi_{\mathrm{ss}}(a,z)$ and $\theta = \theta_{\mathrm{ss}}$ for all (a,z). In that case, the timeless penalty encodes

[29] 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.*"

[30] In Appendix A.8, we also characterize the dual form of the timeless Ramsey problem.

precisely the promises that the Ramsey planner would like to make in the long run, i.e., in the stationary Ramsey plan. Also note that the timeless Ramsey problem nests the standard one, which is time-inconsistent, when we set $\phi(a,z) = 0$ and $\theta = 0$, implying $L^{\text{TP}}(g_0, 0, 0) = L^{\text{SP}}(g_0)$.

A key insight of our approach is that we can transform the standard primal Ramsey problem into a time-consistent, timeless problem by augmenting it with the timeless penalty. Intuitively, by confronting the Ramsey planner with a set of penalties for each forward-looking implementability condition, we introduce an artifical cost that, on the margin, exactly offsets the marginal benefit of time-inconsistent deviations from the stationary Ramsey plan.

Technically, the timeless penalty enforces a set of initial conditions on the promise-keeping multipliers of the primal Ramsey problem. 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 (40) and (39) 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 timeless penalty $\mathcal{T}(\phi, \theta)$, we technically enforce new initial conditions,

$$\phi_0(a,z) = \phi(a,z) \tag{44}$$

$$\theta_0 = \theta, \tag{45}$$

which in turn constrain the planner's choice of initial lifetime values and inflation. Formally, the first-order optimality conditions of the timeless primal Ramsey problem, which define the timeless Ramsey plan, comprise exactly the same equations as in Proposition 4, i.e., equations (33) through (38), except that the initial conditions for the multipliers are now given by (44) and (45). In economic terms, it is for this reason that we refer to $\phi(a,z)$ and θ as *initial promises*. And when we initialize the timeless penalty at $\phi(a,z) = \phi_{\text{ss}}(a,z)$ 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 show that it is time-consistent in Proposition 6, which is the main result of this subsection. At an abstract level, we can interpret all endogenous variables as functions of (i) the policy path, which we denote by $\boldsymbol{i} = \{i_t\}$, the (ii) exogenous shocks, which we denote by $\boldsymbol{Z} = \{A_t, \rho_t, \epsilon_t\}$, and (iii) initial conditions for the distribution $g_0(a,z)$ and promises $(\phi(a,z), \theta)$. This already anticipates the sequence-space representation of our model, which we introduce in Section 5. Hence, given an exogenous sequence of shocks, an initial cross-sectional distribution, as well as initial promises, the planner chooses among different equilibria by choosing a sequence of policy. We can therefore evaluate the timeless Ramsey planner's objective for any given policy path $\boldsymbol{i}$. The timeless Lagrangian then directly implies a local efficiency criterion, according to which the Lagrangian's derivative with respect to the policy path $\boldsymbol{i}$ must be 0 for any locally efficient policy.

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

$$\frac{d}{d\boldsymbol{i}} L^{\text{TP}}\Big(g_{\text{ss}}, \phi_{\text{ss}}, \theta_{\text{ss}}, \boldsymbol{i}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}\Big) = 0. \tag{46}$$

Proposition 6 is an important pillar of our timeless Ramsey approach, and we state its proof in Appendix A.7. Equation (46) 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 timeless penalties of the timeless Lagrangians equal to 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 therefore formally establishes that a planner who sets policy according to the timeless Ramsey problem has no incentive to deviate from the stationary Ramsey plan in the absence of shocks. That is, the timeless Ramsey planner faces no incentive to use surprise inflation either in the long run or in the short run. We illustrate this in Figure 1, which shows that the timeless Ramsey planner always sets inflation to 0 in the absence of shocks.

4.4 Targets for Monetary Policy

We introduced timeless penalties in Section 4.3 and showed that they are critical to address time inconsistency and resolve inflationary bias in the short run. In this subsection, we argue that the timeless penalties θ_{ss} and $\phi_{\text{ss}}(a,z)$ can respectively be interpreted as an *inflation target* in the spirit of Barro and Gordon (1983) and a novel *distributional target*. Connecting to the vast RANK literature on time consistency and targeting rules, we establish three main results: First, an augmented version of the standard inflation target is necessary for monetary policy to be time-consistent in HANK. In the presence of household heterogeneity, the utilitarian planner faces an incentive to use surprise inflation for redistribution. These novel distributional considerations therefore interact with the classical time consistency problem on inflation that emerges from the forward-looking Phillips curve. Second, monetary policy faces a second source of time inconsistency in HANK economies, stemming from households' forward-looking consumption-savings decision, which requires a novel distributional target. We show that the two sources of time inconsistency—aggregate and individual—meaningfully interact now. Third, we justify the interpretation of the timeless penalty as *targets* by showing that policy under discretion, when confronted with the timeless penalty, no longer features inflationary bias in steady state.

4.4.1 Inflation Target

The presence of time inconsistency in the representative-agent New Keynesian model has been the subject of a vast literature starting with Barro and Gordon (1983)—see Clarida et al. (1999) and, in particular, Woodford (2010), who, as we do, works with exact structural equations, for modern treatments.[31] The nature of time inconsistency in RANK models is due to the fact that the planner faces a forward-looking constraint, the Phillips curve, which makes promises over future inflation useful to achieve the planner's objective. In the following, we establish that our timeless penalty on inflation in the RANK limit of our model, i.e., $\theta_{ss}^{RA}\pi_0^{w,RA}$, takes the exact same form as the inflation target derived in the literature that follows Barro and Gordon (1983).[32] The same time consistency problem appears in HANK, where the Phillips curve also incentivizes the planner to make promises about future inflation.

The following Proposition proves an analytical characterization of the inflation target, nesting both our baseline HANK model and its RANK limit, assuming isoelastic preferences with $u(c) = \frac{1}{1-\gamma}c^{1-\gamma}$ and $v(n) = \frac{1}{1+\eta}n^{1+\eta}$.

Proposition 7. (**Inflation Target**) *With isoelastic (CRRA) preferences, the timeless penalty on inflation in both RANK and HANK economies is given by*

$$\theta_{ss} = -\frac{\delta}{\epsilon}\frac{\Omega_{ss}^1 - (\tilde{Y}_{ss}^{RA})^{\gamma+\eta}}{\frac{\epsilon-1}{\epsilon}(1+\tau^L)(1-\gamma)\Omega_{ss}^2 - (1+\eta)(\tilde{Y}_{ss}^{RA})^{\gamma+\eta}}, \tag{47}$$

where $\tilde{Y}_{ss}^{RA}$ is natural output in the RANK limit, defined in Section 2.7, and

$$\Omega_{ss}^1 = 1 + \frac{\iint zu'(c_{ss})\phi_{ss}\,da\,dz}{\iint zu'(c_{ss})g_{ss}\,da\,dz} + \frac{\iint zu''(c_{ss})\chi_{ss}\,da\,dz}{\iint zu'(c_{ss})g_{ss}\,da\,dz}$$

$$\Omega_{ss}^2 = \frac{1}{1-\gamma}\left(1 - \gamma\frac{u'(C_{ss})}{u''(C_{ss})}\frac{\iint z^2u''(c_{ss})g_{ss}\,da\,dz}{\iint zu'(c_{ss})g_{ss}\,da\,dz}\right)$$

are distributional wedges that converge to $\Omega_{ss}^1, \Omega_{ss}^2 \to 1$ in the RANK limit. In the HANK model, $\Omega_{ss}^1, \Omega_{ss}^2 \neq 1$, so distributional considerations shape the inflation target.

The inflation target of Proposition 7 generalizes the standard "Barro-Gordon" inflation target to environments with heterogeneous households. In particular, equation (47) nests the inflation target in both HANK and RANK. The only difference is that the distributional wedges disappear in the RANK limit, i.e., $\Omega_{ss}^1, \Omega_{ss}^2 \to 1$.

In RANK, therefore, the inflation target simplifies to $\theta_{ss}^{RA} = -\frac{\delta}{\epsilon}\frac{1-(\tilde{Y}_{ss}^{RA})^{\gamma+\eta}}{\frac{\epsilon-1}{\epsilon}(1+\tau^L)(1-\gamma)-(1+\eta)(\tilde{Y}_{ss}^{RA})^{\gamma+\eta}}$, which is the standard linear penalty form of the inflation target. It is well understood that time inconsistency only emerges under a distorted steady state so that the planner perceives an incentive to generate surprise inflation and engineer a positive output gap in the short run. In our setting,

[31] It is well understood that RANK economies suffer from a time consistency problem on inflation whenever i) there are cost-push shocks or ii) the appropriate employment subsidy is not in place.

[32] Lemma 33 in the Appendix provides an exact characterization of the marginal benefit a Ramsey planner perceives in the RANK limit from time-inconsistent deviations from the stationary Ramsey plan at time 0. We also prove that the timeless inflation penalty in that limit, $\theta_{ss}^{RA}\pi_0^{w,RA}$, exactly offsets this benefit and can therefore be interpreted as the marginal cost of such time-inconsistent deviations.

this is the case whenever the employment subsidy is not large enough to offset the steady state markup distortion. In fact, it is easy to see that no inflation target is required under the appropriate employment subsidy, with $\frac{\epsilon-1}{\epsilon}(1+\tau^L)=1$, because no time consistency problem emerges in that case.[33]

This is no longer true in HANK. Even with the correct employment subsidy to address the markup distortion in steady state, the time consistency problem on inflation does not disappear. The marginal benefit a Ramsey planner in HANK perceives from time-inconsistent deviations from the stationary Ramsey plan now features a set of distributional motives. Specifically, comparing θ_{ss} in the RANK and HANK models shows that the two sources of time inconsistency in HANK—inflation and redistribution—interact. The planner faces a new redistribution motive for surprise inflation: By overheating the economy with lower nominal interest rates, higher inflation, and a positive output gap, the planner can depress real interest rates in the short run. This benefits indebted, high marginal utility households.

An important corollary of this result is that the choice of the inflation target takes on a distributional dimension whenever the planner has a welfare criterion (or the central bank has a mandate) that goes beyond aggregate efficiency considerations. The standard time consistency problem associated with inflation interacts with the novel redistribution motive of the planner. Crucially, HANK economies feature a time consistency problem on inflation that requires an inflation target even with the correct employment subsidy to address markup distortions in steady state.

4.4.2 Distributional Target

We now characterize the second source of time inconsistency in HANK economies, which is completely absent from RANK economies. It arises because the Ramsey planner disagrees with households on their consumption-savings decisions given the current nominal rate. In particular, a Ramsey planner in a HANK economy finds that households in general save too much or too little, in principle to different degrees. The planner can use the nominal interest rate to adjust consumption-savings decisions on average, and this is sufficient in a RANK economy to fully correct the representative household's consumption-savings decisions. In a HANK economy, however, the planner would still like to modify households' *relative* consumption-savings decisions. Under commitment, therefore, a Ramsey planner finds it optimal to make promises about future interest rates, which in turn manifests via promises on household continuation values. It is through this mechanism that a planner generates promises $\phi_t(a,z)$ over time, opening the door to time inconsistency.

In a RANK economy, it follows from equation (75) that the Ramsey planner sets $\phi_t^{\text{RA}}=0$ at all times. This implies that it is sufficient for the planner to adjust the current nominal interest rate to

[33] To see this, plug in for $\frac{\epsilon-1}{\epsilon}(1+\tau^L)=1$ and set the distributional wedges to 1. Then the steady state with $Y_{ss}^{\text{RA}}=\tilde{Y}_{ss}^{\text{RA}}=1$ implies $\theta_{ss}^{\text{RA}}=0$.

implement any desired level of aggregate consumption-savings. That is, given that the planner can adjust the nominal interest rate, there is no need to further distort the consumption-savings decision of the representative household. And certainly not by making future promises. This eliminates time consistency problems in RANK economies.

In a HANK economy, on the other hand, two equations determine the multiplier $\phi_t(a,z)$. First, the optimality condition for the nominal interest rate, equation (38), implies that a planner seeks to distort individual decisions by setting a weighted average of $\phi_t(a,z)$ equal to

$$\iint a\partial_a \lambda_t(a,z) g_t(a,z)\, da\, dz$$

which captures the *distributive pecuniary effects* of a change in the nominal interest rate. Intuitively, an increase in interest rates benefits savers (with $a > 0$) and hurts borrowers (with $a < 0$). When a planner values dollars in the hands of borrowers and savers differently at the margin—it does so according to the social marginal value of wealth, $\partial_a \lambda_t(a,z)$—this provides a rationale to distort individual decisions.

Second, the optimality condition for household lifetime value $V_t(a,z)$ shows that the evolution of the distribution of individual penalties satisfies the Kolmogorov forward equation (33), which we restate here for convenience:

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

Intuitively, this KF equation implies that the distribution of penalties must be consistent with the law of motion of households across the different individual states, summarized by the operator $\mathcal{A}_t^*$.[34] The term $\partial_a \chi_t(a,z)$ is a forcing term and thus a source of time variation in the distribution of promises $\phi_t(a,z)$. If individuals' optimal consumption and savings decisions change as they transition between different wealth states, then $\partial_a \chi_t(a,z)$ can be interpreted as "births" and "deaths" of relative promises in the cross section. Solving for ϕ_t and χ_t jointly and characterizing how they are linked via the Kolmogorov forward equation (33) is one of the contributions of this paper.[35]

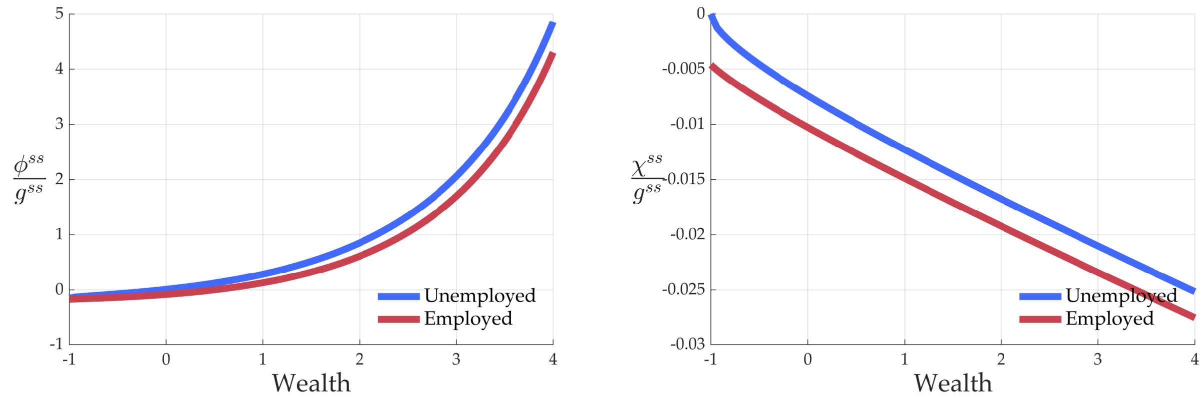
Proposition 8. (Distributional Target)

a) A utilitarian planner faces a distributional time consistency problem, with distributional penalties that evolve via the Kolmogorov forward equation defined in equation (33).

b) An Aggregate Efficiency planner—see Dávila and Schaab (2021)*—does not face a distributional time consistency problem.*

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

[35] It should be evident that if households made multiple individual decisions in addition to the consumption and savings decisions, these would enter equation (33) as additional forcing terms.



(a) Individual penalties **(b)** Savings multiplier

Figure 2. Distributional Target

Note. The left panel of Figure 2 shows the steady state values of the individual target, $\phi_{\text{ss}}(a,z)$, normalized by the mass of households, $g_{\text{ss}}(a,z)$. The right panel of Figure 2 shows the steady state values of the savings multiplier, $\chi_{\text{ss}}(a,z)$, also normalized by the mass of households, $g_{\text{ss}}(a,z)$. From the right panel, it immediately follows that the planner perceives that all agents are under-saving. Note that we formally show that $\iint \phi_{\text{ss}}(a,z)\, g_{\text{ss}}(a,z)\, da\, dz = 0$. Note that the planner perceives that a household is under-saving when $\chi_{\text{ss}}(a,z) < 0$.

An important conclusion from this section is that if a planner (central bank) adopts a welfare criterion (mandate) that is not the aggregate efficiency one, and if the planner (central bank) would like to implement a time-consistent policy, then it must also adopt what we call a *distributional target* in addition to the standard inflation target. In other words, outside of the aggregate efficiency criterion, the distributional target becomes necessary for optimal policy.

Finally, it may be helpful to illustrate the form of the distributional penalties, as we do in Figure 2. The left panel of Figure 2 plots the distributional penalties associated with the stationary Ramsey plan, $\phi_{\text{ss}}(a,z)$, normalized by the mass of households, $g_{\text{ss}}(a,z)$. The right panel of Figure 2 displays the stationary consumption-savings multiplier, $\chi_{\text{ss}}(a,z)$, also normalized by the mass of households, $g_{\text{ss}}(a,z)$. From the right panel, it immediately follows that the planner perceives that all households are under-saving and over-consuming, since $\chi_{\text{ss}}(a,z) < 0$, with poorer and unemployed households less so.

4.4.3 Targets, Discretion, and Inflationary Bias

We conclude our discussion of targets by briefly revisiting the policy problem under discretion. Consider again problem (20), where a Ramsey planner sets policy with commitment over a finite horizon. Our discussion in Section 3 focuses on the limit of this problem as the commitment horizon becomes vanishingly small, i.e., as $\psi \to \infty$, which we refer to as policy under discretion. We now revisit this discussion in the context of timeless penalties and targets. We show—in the spirit of the canonical analysis of *targets* (Galí, 2015)—that timeless penalties correspond exactly to the policy target—or penalty, in the language of a transferable utility mechanism—that aligns (Markov) policy under discretion with policy under commitment in steady state.

Define the timeless Ramsey problem with finite commitment horizon by

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

where $\boldsymbol{M}$, $\boldsymbol{X}$, and $L(0, T, g_0)$ are defined as in (20). In words, we confront *each* finite-horizon Ramsey planner with the timeless penalties $\mathcal{T}(\phi, \theta)$. Consider again the limit $\psi \to \infty$, which we continue to associate with policy under discretion. In this limit, Ramsey planners transition instantaneously and their commitment horizon becomes vanishingly small. But whenever a new Ramsey planner takes over, she is again confronted with the timeless penalties. In the limit, therefore, these penalties are active in every instant. In the discrete-time analysis of, e.g., Galí (2015), we would say that the planner faces these penalties in every period.[36] In this sense, the resulting sequence of timeless penalties $\mathcal{T}(\cdot)$ takes the form of a linear penalty *target* using the language of the mechanism design literature following Barro and Gordon (1983) and Walsh (1995). We now obtain the following result.

Proposition 9. (**Monetary Policy under Discretion with a Target**) *The Markov perfect stationary equilibrium that corresponds to problem* (48) *as* $\psi \to \infty$*, which we refer to as* policy under discretion with targets*, features no inflationary bias. That is,* $\pi_{ss} = 0$*.*

Proposition 9 underscores again that the timeless penalties we introduced and characterized in this section should be thought of as targets: an augmented inflation target and a new distributional target. The resulting problem (48) resembles the kind of transferable utility mechanism that has been shown to address time inconsistency in the standard New Keynesian model (Walsh, 1995). Proposition 9 shows that our timeless penalty aligns monetary policy in the Markov perfect stationary equilibrium under discretion with optimal policy under the stationary Ramsey plan.

4.5 Optimal Stabilization Policy

Finally, we compare optimal monetary stabilization policy in HANK and RANK. In the classical representative-agent analysis, prescriptions for monetary stabilization policy have commonly been summarized in the form of *targeting rules* (Clarida et al., 1999; Galí, 2015). We follow this tradition in our analysis. In fact, our main result in this subsection is a non-linear, i.e., exact, targeting rule for optimal monetary stabilization policy in response to demand, productivity, and cost-push shocks that nests both RANK and HANK. As in our discussion around equation (19), we denote by $\tilde{Y}_t^{\text{RA}}$

[36] Equation (48) closely resembles the optimal transferable utility mechanism that emerges in this kind of environment. See for example Clayton and Schaab (2021) for a recent treatment. Establishing formally whether our timeless penalty corresponds to the optimal transferable utility mechanism in our environment goes beyond the scope of this paper.

natural output in the RANK limit of our economy, which varies with productivity and cost-push shocks, but not with demand shocks.

Proposition 10. (**Targeting Rule for Stabilization Policy under Commitment**) *We summarize optimal monetary stabilization policy with the targeting rule*

$$Y_t = \tilde{Y}_t^{RA} \left\{ \frac{\frac{\epsilon_t}{\epsilon_t - 1}(1+\tau^L)\Omega_t^1 + \theta_t(1-\gamma)\frac{\epsilon_t}{\delta}\Omega_t^2}{1+\theta_t(1+\eta)\frac{\epsilon_t}{\delta}} \right\}^{\frac{1}{\gamma+\eta}}. \tag{49}$$

We start by revisiting the classical results on monetary stabilization policy in RANK through the lens of targeting rule (49). In RANK, we have $\Omega_t^1 = \Omega_t^2 = 1$. Suppose that we also allow for the appropriate steady state employment subsidy, so that $\frac{\epsilon-1}{\epsilon}(1+\tau^L) = 1$ and there is no steady state distortion due to monopolistic competition.

Consider first the case of demand and TFP shocks. Suppose the planner closes the wage inflation gap so that $\pi_t^{w,\text{RA}} = 0$. This implies $\theta_t^{\text{RA}} = 0$ from equation (73) because, under zero inflation, the planner never builds up inflation promises. Consequently, the term in brackets in equation (49) simply becomes 1, and we have $Y_t^{\text{RA}} = \tilde{Y}_t^{\text{RA}}$. In response to demand and TFP shocks, the Ramsey planner in RANK wants to close both the inflation and output gap. This is the seminal Divine 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. Suppose again that the planner closes the wage inflation gap, implying $\pi_t^{w,\text{RA}} = \theta_t^{\text{RA}} = 0$. With $\epsilon_t \neq \epsilon$, we have $Y_t^{\text{RA}} = \tilde{Y}_t^{\text{RA}} \{\frac{\epsilon_t}{\epsilon_t - 1}(1+\tau^L)\}^{\frac{1}{\gamma+\eta}}$. The planner consequently does not find it optimal to close the inflation and output gaps at the same time.

In a HANK economy, we generically have $\Omega_t^1, \Omega_t^2 \neq 1$. In the presence of these distributional wedges, Divine Coincidence fails even with the appropriate employment subsidy: the Ramsey planner never finds it optimal to close both output and inflation gaps at the same time. In other words, the Ramsey planner always perceives a tradeoff between aggregate stabilization, i.e., the inflation and output gaps, on the one hand, and distributional considerations on the other hand.

At the same time, Proposition 10 tells us that accounting for distributional considerations in stabilization policy generically comes at the cost of aggregate efficiency. For a planner (central bank mandate) that only values aggregate efficiency, Divine Coincidence is restored (Dávila and Schaab, 2021). The targeting rule for optimal monetary stabilization policy for such a planner is the same as the RANK targeting rule.

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—with and without commitment—efficiently and fast. In particular, we show that the timeless Ramsey approach we developed in Section 4 allows us to leverage sequence-space perturbation methods and extend the fake-news algorithm of (Auclert et al., 2021) to compute optimal policy dynamics.

5.1 Sequence Space Methods for Optimal Policy in HANK

To characterize optimal stabilization policy, we work with an abstract sequence-space representation of our model. The results we develop in this section build on recent work by Boppart et al. (2018) and Auclert et al. (2021), and extend the sequence-space approach to Ramsey problems and welfare analysis in heterogeneous-agent economies.

Equilibria in our baseline HANK economy can be summarized by an *equilibrium map* that take as inputs the time paths of aggregates,

$$\mathcal{H}(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z}) = 0, \tag{50}$$

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 (50) 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 (50) of general equilibrium in our model is as in 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 11. (**Sequence-Space Representation of Ramsey Plans**) *Given an initial distribution $g_0(a,z)$, initial promises $\phi(a,z)$ and θ, as well as a path for exogenous shocks $\boldsymbol{Z}$, a Ramsey plan jointly characterizes macroeconomic allocations and prices $\boldsymbol{X}$, optimal policy $\boldsymbol{\theta}$, and multipliers $\boldsymbol{M}$. Its sequence-space representation is given by*

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

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 (50) and Ramsey plans (51) in Appendices B.1 and B.2.

Our sequence-space representation of Ramsey plans in heterogeneous-agent economies is valid for any initial distribution $g_0(a,z)$ and initial promises $\phi(a,z)$ and θ. Equation (51) 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 directly from Proposition 6 that the Ramsey plan characterized by (51) is time-consistent when instead evaluated at $(g_{\text{ss}}, \phi_{\text{ss}}, \theta_{\text{ss}})$. In that case, we refer to it as a timeless Ramsey plan. In the following, we always initialize the timeless penalty 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 (51) consists of two sets of equations. The first block is the system of equations (50), 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.[37]

In this sequence-space representation, we refer to a (timeless) Ramsey plan as the time paths of aggregates, $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i})$.[38] The system of equations (51) characterizes a Ramsey plan as a 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 promises $\phi(a,z)$ and θ. The sequence-space representation of Ramsey plans $\boldsymbol{R}$ in Proposition 11 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 (51) becomes a system of five equations: the definition of Λ_t as the aggregate labor wedge, the resource constraint (18), as well as the three aggregate optimality conditions (36), (37), and (38). 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 promises $\phi(a,z)$ and θ.

Finally, notice that while Proposition 11 develops a sequence-space representation of Ramsey plans in heterogeneous-agent economies, it is a non-linear, i.e., exact, representation. We will not make use of sequence-space perturbation methods, i.e., approximate representations, until Section 5.1.2.

[37] In this sense, Proposition 11 and the sequence-space representation of Ramsey plans (51) are closely associated with the timeless primal Ramsey problem, rather than its dual form. This discussion already anticipates that solving the Ramsey plan in its primal sequence-space representation requires solving for the transition dynamics of multipliers, as we further discuss below.

[38] 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 B.1 and B.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.

5.1.1 Non-Linear Optimal Policy

The sequence-space representation of a timeless Ramsey plan in our economy is a system of non-linear equations. We can directly solve (51) 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 timeless Ramsey plans non-linearly is tractable and fast in our baseline HANK economy. Using an efficient quasi-Newton algorithm, we can solve the system (51) non-linearly in less than 10 seconds.[39] 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.[40]

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.

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 (51).

Proposition 12. (**Optimal Policy Perturbations in the Primal**) *Consider the primal Ramsey problem and the associated Ramsey plan, which is characterized by (51) and solves $\mathcal{R}(\cdot) = 0$. Suppose we initialize the Ramsey plan at the cross-sectional distribution $g_0(a,z) = g_{\text{ss}}(a,z)$ and with initial timeless penalties $\phi(a,z) = \phi_{\text{ss}}(a,z)$ and $\theta = \theta_{\text{ss}}$. To first order, optimal stabilization policy is then characterized as part of the Ramsey plan by*

$$d\boldsymbol{R} = -\mathcal{R}_{\boldsymbol{R}}^{-1}\mathcal{R}_{\boldsymbol{Z}}d\boldsymbol{Z} \tag{52}$$

where $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{\text{ss}}$ 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.

[39] We use the quasi-Newton algorithm developed by Schaab and Zhang (2021) 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 non-linear Ramsey plans for our baseline HANK economy takes less than 10 seconds on a personal computer for discretized time grids with up to 150 nodes. For this experiment, we use a 2020 13-inch MacBook Pro with an M1 chip and 16 GB memory.

[40] 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 $\mathcal{R}_{\boldsymbol{Z}}$.

We prove Proposition 12 in Appendix B.3.

Note that the characterization of stabilization policy in Proposition 12 relies on initializing the Ramsey problem with the proper timeless penalties, so that $\mathcal{R}(\cdot)$ characterizes a timeless Ramsey plan. With the appropriate 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 time consistency problems 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 B, we extend the fake-news algorithm developed by Auclert et al. (2021) for sequence-space Jacobians to compute optimal policy and Ramsey plans.[41]

5.1.3 Optimal Policy Perturbations in the Dual

Appendix A.8 formally introduces the dual form of our timeless Ramsey problem. While the previous subsection directly leverages the primal representation of Ramsey plans, an alternative sequence-space perturbation method can be developed by using the dual form as a starting point. We do so in Appendix A.9.

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 B, we therefore introduce *sequence-space Hessians* as the natural second-order generalization of sequence-space Jacobians. Finally, Appendix A.9 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 Experiments

In this section, we compute and characterize optimal monetary stabilization policy in response to different types of shocks, comparing results for our HANK baseline economy against the classical RANK benchmark. The primary goal of our quantitative analysis is to further shed light on the key

[41] Auclert et al. (2021) show how to use the equilibrium map (50) 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 B that computing optimal policy using the sequence-space Ramsey plan representation (51) 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.

departures from the classical consensus on optimal monetary policy as, for example, summarized in Clarida et al. (1999), Woodford (2003), and Galí (2015).

Calibration. We adopt isoelastic preferences and use standard parameters for household preferences in both model benchmarks, setting 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).

The main difference between our representative- and heterogeneous-agent benchmarks is the earnings process that households face. In our HANK model, we assume that z_t follows a two-state Markov chain, with $z_t \in \{z^{\mathrm{L}}, z^{\mathrm{H}}\}$, where $z^{\mathrm{L}} = 0.8$ and $z^{\mathrm{H}} = 1.2$. We set the quarterly Poisson transition rate out of both states to 0.33. Our RANK benchmark can be seen as the limit as $z^{\mathrm{L}}, z^{\mathrm{H}} \to \bar{z} = 1$, and using as initial condition for the cross-sectional distribution a Dirac mass at $(a, z) = (0, \bar{z})$.

Finally, in the following exercises we assume that the planner sets the employment subsidy τ^L to address the wage-markup distortion in stationary equilibrium. That is, $(1 + \tau^L)\frac{\epsilon - 1}{\epsilon} = 1$. Since much of the analysis of monetary policy in RANK has made use of such an employment subsidy, we also report results in the main text under this assumption to facilitate comparison to the classical results.

Shocks. Following the standard New Keynesian literature, we consider three types of shocks—demand, productivity, and cost-push shocks. We assume that the three shocks follow mean-reverting AR(1) processes. In continuous time, this implies that

$$\dot{A}_t = \xi_A (A - A_t), \quad \dot{\epsilon}_t = \xi_\epsilon (\epsilon - \epsilon_t), \quad \text{and} \;\; \dot{\rho}_t = \xi_\rho (\rho - \rho_t)$$

where A, ϵ, and ρ denote the steady-state constant levels. We study one-time, unanticipated ("MIT") shocks at time $t = 0$, so that A_0, ϵ_0, and ρ_0 jump and subsequently revert back to their steady-state levels following the above laws of motion. We set the initial shock levels to $A_0 = 1.005A$, $\epsilon_0 = 1.25\epsilon$, and $\rho_0 = 1.5\rho$, and calibrate the shock's persistence in each case to a half-life of one quarter.

Transition dynamics with interest rate rules. For reference, we also compute the transition dynamics of both HANK and RANK models for the case where monetary policy is not set optimally but instead follows a standard interest rate rule. We report these results in Appendix F.3.

Sensitivity analysis. We perform additional sensitivity and robustness analysis in Appendix F.4. In particular, we replicate the following numerical experiments for economies featuring differing degrees of earnings and consumption inequality. These exercises underscore that uninsurable

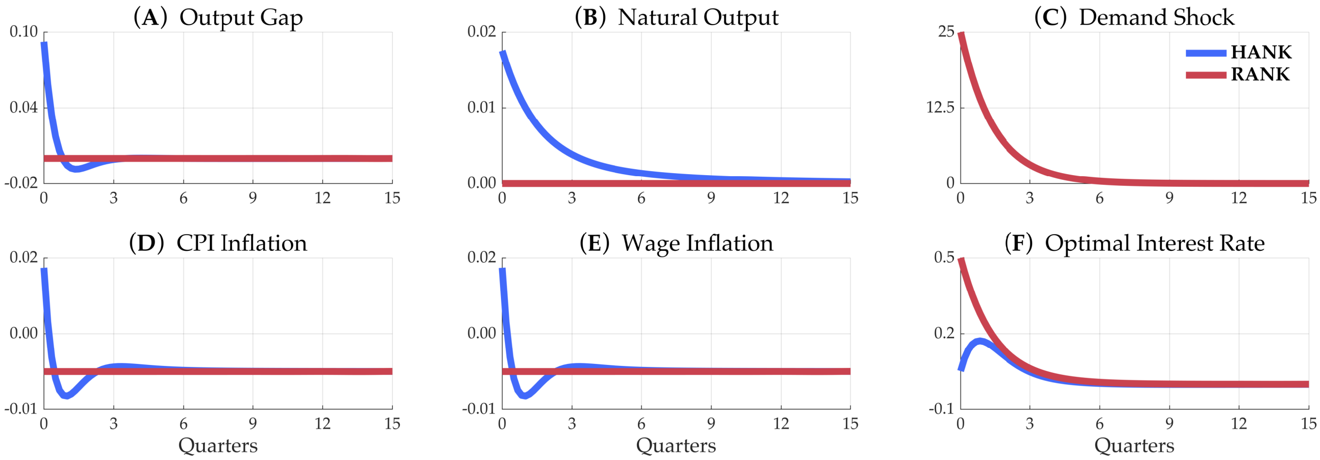


Figure 3. Optimal Policy Transition Dynamics: Demand Shock

Note. Transition dynamics after a positive discount rate shock in both RANK (red) and HANK (blue) models under optimal monetary stabilization policy. 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.

earnings risk and incomplete markets are the key economic force that motivates departures of optimal policy from the classical targeting rules in RANK.

5.2.1 Demand Shocks

We start by discussing optimal stabilization policy in response to a demand shock. We display the associated transition dynamics in Figure 3. See Figure 10 in Appendix F.3 for the analogous impulse responses under a Taylor rule. Demand shocks have no impact on natural output in RANK. In HANK, we see natural output increase slightly after a positive demand shock because unions' valuation of wages changes.

The classical result on optimal monetary stabilization policy in RANK is that Divine Coincidence obtains in the face of demand and productivity shocks: the planner perfectly stabilizes both the output and inflation gaps. This sharp benchmark result requires the appropriate employment subsidy, of course, which we assume here. To support this desired allocation, the 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

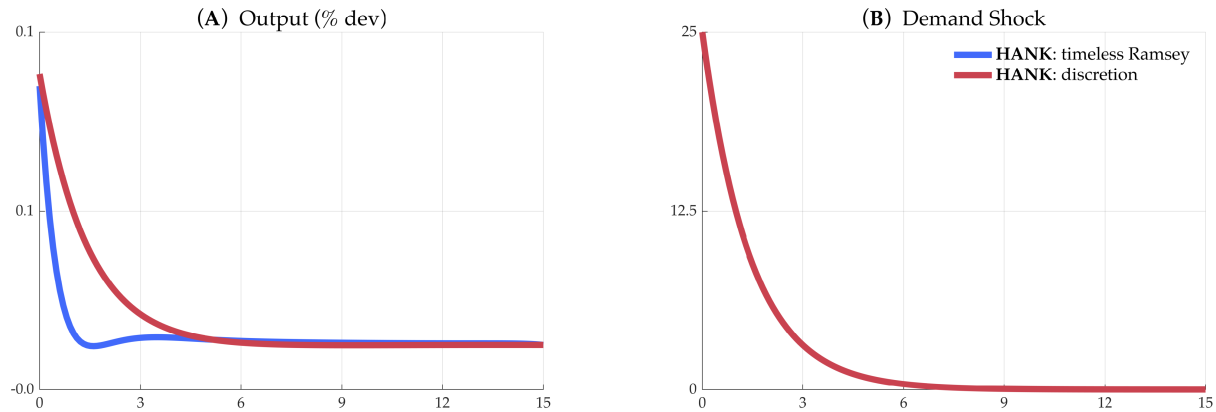


Figure 4. Optimal Policy under Discretion: Demand Shock

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.

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 under optimal monetary policy in Figure 3 are due to commitment. In Figure 4, we plot the relative output paths when optimal monetary policy is set under discretion (red) and under the timeless Ramsey problem (blue). Under discretion, the planner cannot manipulate expectations about future policy and has no incentive to promise an over-shooting. With commitment, the planner finds it optimal to promise over-shooting of both output and inflation. This is reflected in the hump-shaped optimal interest rate path in Figure 3.

5.2.2 Productivity Shocks

We next turn to optimal stabilization policy in response to a TFP shock. Figure 5 reports the transition dynamics of the economy under optimal policy, while Figure 9 in Appendix F.3 reports those under a Taylor rule for comparison.

In both model benchmarks, natural output increases in response to a positive productivity shock. Natural output increases less than one-for-one, primarily due to diminishing marginal utility from consumption and convex disutility from labor (see Section 2.7). In HANK, natural output increases slightly less than in RANK as a result of union wage bargaining, which now features a distributional consideration.

Optimal stabilization policy in HANK follows the same principles as in RANK, with minor

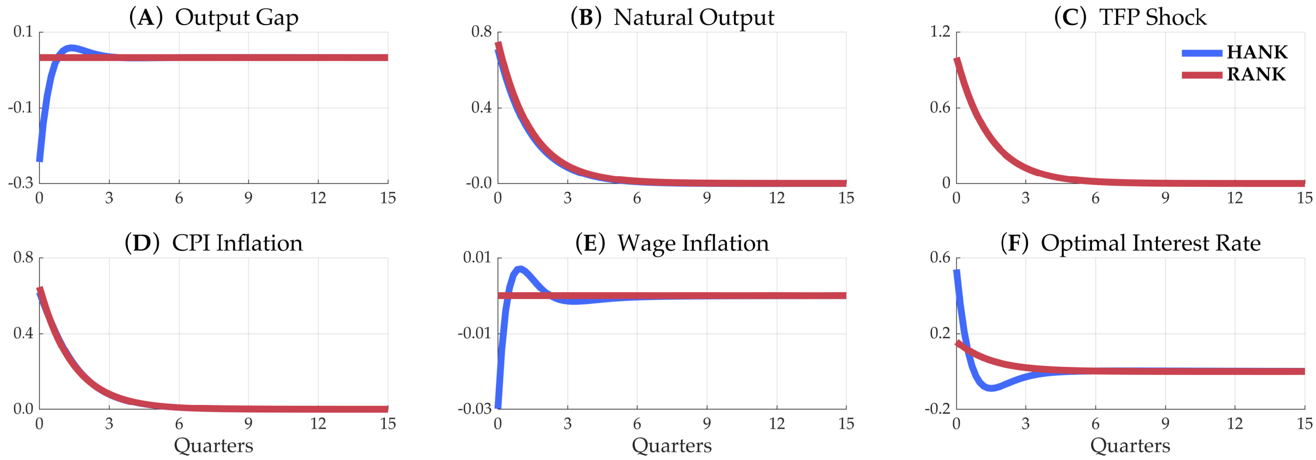


Figure 5. Optimal Policy Transition Dynamics: TFP Shock

Note. Transition dynamics after a positive TFP shock in both RANK (red) and HANK (blue) models under optimal monetary stabilization policy. Initial shock is 1% of steady state TFP and mean-reverts 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.

quantitative departures. The planner largely stabilizes both output and (wage) inflation gaps, but not fully. The planner allows both to become briefly negative on impact, before becoming positive and overshooting, yielding a hump-shaped response. The wage inflation gap on impact is small, reaching only -0.02%, and consequently not meaningfully different from 0. Compared to the response of wage inflation under a Taylor rule, where the wage inflation gap opens up to 0.4% under the same shock, this deviation from the Divine Coincidence benchmark of RANK should be viewed as minimal. Similarly, while optimal policy stabilizes the output gap substantially relative to policy under the Taylor rule, the planner allows a small negative output gap to open up. The on-impact negative output gap under optimal policy is less than 20% of the size of the output gap under the Taylor rule.

5.2.3 Cost-Push Shocks

Finally, we consider a cost-push shock under which the desired wage mark-up of labor unions changes and natural output increases by 0.25%. We report the transition dynamics under optimal policy in Figure 6, and also report the analogous transition dynamics under a Taylor rule in Figure 11 in Appendix F for comparison.

In RANK, Divine Coincidence fails in the presence of cost-push shocks and the planner now faces a tradeoff between inflation and output. Optimal stabilization policy is accommodative, lowering the nominal interest rate, but a small negative output gap still opens up.

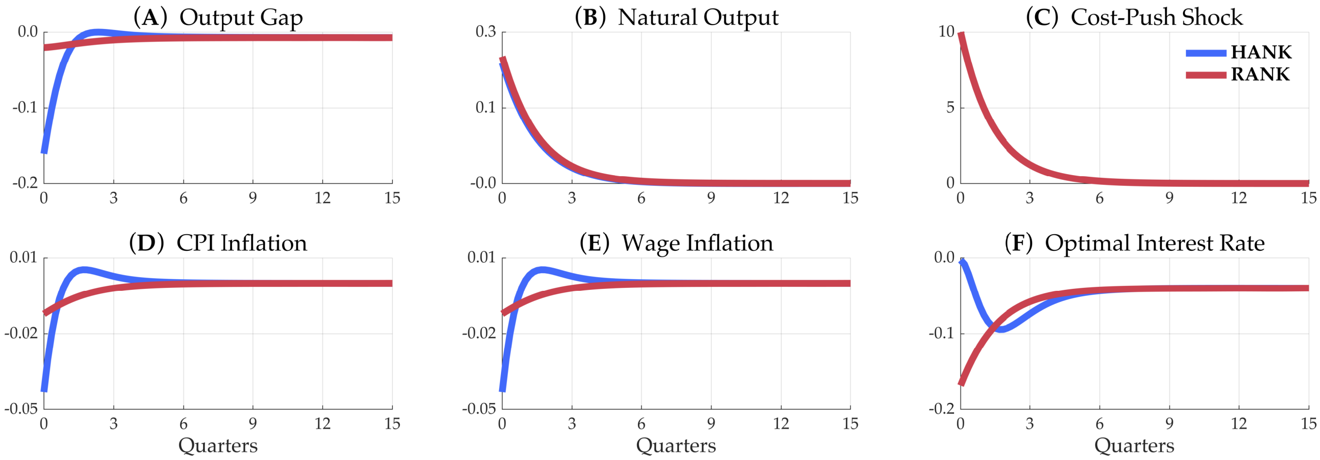


Figure 6. Optimal Policy Transition Dynamics: Cost-Push Shock

Note. Transition dynamics after a positive cost-push shock in both RANK (red) and HANK (blue) models under optimal monetary stabilization policy. The cost-push shock is modeled as an increase in labor union's desired wage mark-up. The shock is initialized at $\epsilon_0 = 11$ and mean-reverts to its steady state value $\epsilon = 10$, with a half-life of 1 quarter. Panels (**A**) through (**C**) report the dynamics of the output gap, $\frac{Y_t - \bar{Y}_t}{\bar{Y}_t}$, 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.

In HANK, natural output again increase but slightly less due to distributional concerns in union bargaining. Monetary policy eases substantially less than in RANK, allowing a sizable negative output gap to open up. However, there is still substantial stabilization relative to the Taylor rule case. Especially inflation is again stabilized substantially.

6 Conclusion

In this paper, we have characterized optimal monetary policy in a HANK economy in which households face uninsurable idiosyncratic risk and wages are rigid. We showed that policy under discretion is subject to a novel redistribution motive, which leads to exacerbated inflationary bias in steady state. Following a timeless Ramsey approach, we have studied optimal long-run policy and optimal stabilization policy under commitment, analyzed the question of time consistency, and analyzed optimal policy under discretion. We show i) that zero inflation is the optimal long-run policy (in our baseline model), ii) that a planner faces two sources of time inconsistency (inflation and distributional) that non-trivially interact with each other, and iii) that the Divine Coincidence does not hold in HANK economies in response to non-cost-push shocks, even when it would in RANK economies.

Lastly, our paper establishes that time-consistent, optimal monetary stabilization policy generically requires a distributional target—alongside the classical inflation target—whenever a welfare criterion is used that has distributional considerations beyond aggregate efficiency. If societies consider expanding central banks' mandates in the future to include distributional considerations, they will also confront a new set of time consistency problems that may warrant a revision of the classical targeting framework.

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. 2021. Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach.

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. 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.

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

Bianchi, Javier and Enrique G Mendoza. 2018. Optimal Time-Consistent Macroprudential Policy. *Journal of Political Economy*, 126(2):588–634.

Bilbiie, Florin O. 2020. The new Keynesian cross. *Journal of Monetary Economics*, 114:90–108.

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.

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, Xavier Jaravel, and Andreas Schaab. 2018. Heterogeneous price rigidities and monetary policy. *Available at SSRN 3186438*.

Clayton, Christopher and Andreas Schaab. 2021. 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.

Cravino, Javier, Ting Lan, and Andrei A Levchenko. 2020. Price stickiness along the income distribution and the effects of monetary policy. *Journal of Monetary Economics*, 110:19–32.

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. 2021. Welfare Assessments with Heterogeneous Individuals.

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.

Jeanne, Olivier and Anton Korinek. 2020. Macroprudential regulation versus mopping up after the crash. *The Review of Economic Studies*, 87(3):1470–1497.

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.

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.

Schaab, Andreas. 2020. Micro and Macro Uncertainty.

Schaab, Andreas and Allen Tianlun Zhang. 2021. Dynamic Programming in Continuous Time with Adaptive Sparse Grids.

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

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

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

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.

A Proofs and Derivations for Section 4

In this Appendix, we present the proofs and main derivations for the results of Section 4. We start in Appendix A.1 by stating formally the implementability conditions for the Ramsey problem of Section 4 in continuous time. We also provide additional details on the generator and adjoing, $\mathcal{A}_t$ and $\mathcal{A}_t^*$, which we use in the main text.

We then formally state and solve the standard primal Ramsey problem in Appendix A.2. The resulting optimality conditions, which we report in the main text in Proposition 4 hold everywhere in the interior of the state space over (a, z).

A key challenge in solving Ramsey problems with heterogeneous agents is to formally account for boundary conditions, in particular the borrowing constraint at $\underline{a}$. We find it convenient to derive all proofs that explicitly account for the boundary of the state space in a discretized version of our model. We follow Achdou et al. (2021) and work with a consistent finite-difference discretization of our continuous-time heterogeneous-agent equations, which of course converge in the limit to our baseline HANK economy. We follow this approach in the remainder of this appendix.

In recent work, González et al. (2021) follow a similar approach, first casting the optimal policy problem in continuous time, and then discretizing the resulting Ramsey plan conditions. The main difference between our paper and theirs is that they directly take their discretized system of equations to Dynare to obtain a numerical characterization of the Ramsey plan. We leverage the discretized equations to prove the main results on time consistency for our timeless Ramsey approach in Section 4. Our primary interest in discretizing the Ramsey plan conditions is to properly take into account the borrowing constraint faced by households, as well as the distribution mass point that emerges at the borrowing constraint.

A.1 Competitive Equilibrium and Implementability

A competitive equilibrium of our baseline HANK model can be characterized by three blocks of equations. First, there is an individual block, explained in the text, which corresponds to the households' HJB, their optimality condition for consumption, and the Kolmogorov forward equation:

$$\rho V_t(a,z) = \partial_t V_t(a,z) + u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 + \mathcal{A}_t V_t(a,z),$$

$$u'(c_t(a,z)) = \partial_a V_t(a,z)$$

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

$\mathcal{A}_t$ is the infinitesimal generator of the process (a_t, z_t). Intuitively, it captures an agent's perceived law of motion of the process $d(a_t, z_t)$. It is analogous to a transition matrix in discrete time, and it is defined by

$$\mathcal{A}_t f_t(a,z) = \Big(r_t a + z w_t N_t - c_t(a,z)\Big)\partial_a f_t(a,z) + \mathcal{A}_z f_t(a,z), \tag{53}$$

for any function $f_t(a,z) : \mathbb{R}^3 \to \mathbb{R}$, where $\mathcal{A}_z$ is an additively separable component that captures perceived transition dynamics of earnings risk. We leave the structure of $\mathcal{A}_z$ fully general in our derivations, except that we assume it to be independent from policy. Our baseline results currently do not apply to the case of counter-cyclical earnings risk that responds to monetary policy, for example, but extending our approach to this more general case is straightforward.

We denote the adjoint of the infinitesimal generator by $\mathcal{A}_t^*$. The adjoint is defined by

$$\mathcal{A}_t^* f_t(a,z) = -\partial_a\Big[\Big(r_t a + z w_t N_t - c_t(a,z)\Big) f_t(a,z)\Big] + \mathcal{A}_z^* f_t(a,z), \tag{54}$$

where $\mathcal{A}_z^*$ is the adjoint of $\mathcal{A}_z$.

Second, there is an aggregate block, which includes the New Keynesian wage Phillips curve, the production technology, the wage equation, the Fisher equation, and and equation that relates price and wage inflation:

$$\begin{aligned}
\dot{\pi}_t^w &= \rho \pi_t^w + \frac{\epsilon}{\delta}\left[\frac{\epsilon - 1}{\epsilon}(1+\tau^L) w_t \Lambda_t - v'(N_t)\right] N_t \\
Y_t &= A_t N_t \\
w_t &= A_t \\
r_t &= i_t - \pi_t \\
\pi_t &= \pi_t^w - \frac{\dot{A}_t}{A_t}.
\end{aligned}$$

Finally, we have the market clearing conditions in the goods and bond markets, given by

$$\begin{aligned}
Y_t = C_t &= \iint c_t(a,z)\, g_t(a,z)\, da\, dz \\
0 = B_t &= \iint a g_t(a,z)\, da\, dz.
\end{aligned}$$

The following Lemma defines the set of implementability conditions that act as constraints for a Ramsey planner.

Lemma 13. (**Implementability conditions**) *The set of equations that define an equilibrium can be* *expressed as implementability conditions for a standard primal Ramsey problem as follows:*

$$\begin{aligned}
\rho V_t(a,z) &= \partial_t V_t(a,z) + u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 + \mathcal{A}_t V_t(a,z) \\
u'(c_t(a,z)) &= \partial_a V_t(a,z) \\
\partial_t g_t(a,z) &= \mathcal{A}_t^* g_t(a,z) \\
0 &= A_t N_t - \iint c_t(a,z) g_t(a,z)\, da\, dz \\
r_t &= i_t - \pi_t^w + \frac{\dot{A}_t}{A_t} \\
\dot{\pi}_t^w &= \rho \pi_t^w + \frac{\epsilon}{\delta}\left[\frac{\epsilon - 1}{\epsilon}(1+\tau^L) A_t \iint z u'(c_t(a,z)) g_t(a,z) da dz - v'(N_t)\right] N_t
\end{aligned}$$

A.2 Derivation of Ramsey Plan in Continuous Time

In this section, we prove Proposition 4 and derive the continuous-time optimality conditions that characterize the optimal Ramsey plan in the interior of the state space. We defer a formal treatment of the boundary and associated boundary conditions to the following sections.

It is useful to adopt more compact notation for this derivation. In particular, we drop time subscripts and make implicit the dependence of individual variables on states, so that $c_t(a,z)$ simply becomes c. Furthermore, we now reserve subscripts to denote partial derivatives, so that $\partial_t c_t(a,z)$ will simply become c_t.

The functional Lagrangian associated with the standard primal Ramsey problem is given by

$$\begin{aligned}
L^{\mathrm{SP}}(g_0) = \int_0^\infty e^{-\rho t} \Bigg\{ \iint \Bigg\{ &\left[u(c) - v(N) - \frac{\delta}{2}(\pi^w)^2\right] g \\
&+ \phi \left[-\rho V + V_t + u(c) - v(N) - \frac{\delta}{2}(\pi^w)^2 + \mathcal{A} V \right] \\
&+ \chi \left[u'(c) - V_a \right] \\
&+ \lambda \left[-g_t + \mathcal{A}^* g \right] \Bigg\} da dz \\
&- \mu \left[\int \int c g da dz - AN \right] \\
&+ \vartheta \left[-\dot{\pi}^w + \rho \pi^w + \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L) A \Lambda - v'(N) \right) N \right] \Bigg\} dt
\end{aligned}$$

We now use the following auxilliary results, integrating various partial derivatives in the above Lagrangian by parts. We have

$$\int_0^\infty \iint \Big[e^{-\rho t}\phi V_t\Big] d(a,z)dt = \int \Big[-\nu(0,a,z)V(0,a,z) + \rho\int_0^\infty e^{-\rho t}\phi V dt - \int_0^\infty e^{-\rho t}\phi_t V dt\Big] d(a,z)$$

$$\int_0^\infty \iint \Big[e^{-\rho t}\lambda g_t\Big] d(a,z)dt = \int \Big[-\lambda(0,a,z)g(0,a,z) + \rho\int_0^\infty e^{-\rho t}\lambda g dt - \int_0^\infty e^{-\rho t}\lambda_t g dt\Big] d(a,z)$$

$$\int_0^\infty \Big[e^{-\rho t}\vartheta\pi_t\Big] dt = -\vartheta(0)\pi(0) + \rho\int_0^\infty e^{-\rho t}\vartheta\pi dt - \int_0^\infty e^{-\rho t}\vartheta_t\pi dt.$$

Next, for the adjoint, we have

$$-\int_0^\infty e^{-\rho t}\iint \lambda\mathcal{A}^* g d(a,z)dt = -\int_0^\infty e^{-\rho t}\iint (\mathcal{A}\lambda) g d(a,z)dt,$$

where we drop boundary terms, which we consider formally in the following subsections. And for the generator, we have

$$\int_0^\infty e^{-\rho t}\int \phi\mathcal{A}V dadzdt = \int_0^\infty e^{-\rho t}\iint V\mathcal{A}^*\phi dadzdt$$

Finally, for the consumption FOC, we simply have

$$-\int_0^\infty e^{-\rho t}\iint \chi V_a dadzdt = \int_0^\infty e^{-\rho t}\iint \chi_a V dadzdt,$$

where we also drop boundary terms.

The functional Lagrangian can thus be rewritten as

$$\begin{aligned} L^{\text{SP}}(g_0) = \int_0^\infty e^{-\rho t}\Bigg\{ \iint \Bigg\{ & \Big[u(c) - \mu c - v(N) - \frac{\delta}{2}(\pi^w)^2\Big] g \\ & - V\phi_t + V\mathcal{A}^*\phi + \phi\Big[u(c) - v(N) - \frac{\delta}{2}(\pi^w)^2\Big] \\ & + \chi u'(c) + \chi_a V \\ & + g\lambda_t - \rho\lambda g + g\mathcal{A}\lambda \Bigg\} dadz \\ & + \mu AN \\ & + \vartheta_t\pi^w + \vartheta\frac{\epsilon}{\delta}\Big(\frac{\epsilon-1}{\epsilon}(1+\tau^L)A\Lambda - v'(N)\Big)N \Bigg\} dt \end{aligned}$$

We now consider a general functional perturbation around a candidate optimal Ramsey plan, and parametrize this perturbation by $\alpha \in \mathbb{R}$. Since α is a scalar, the maximum principle then implies that our candidate plan can only be optimal if $L^{\mathrm{SP}}_\alpha(g_0, \alpha) \mid_{\alpha=0} = 0$.

We have

$$
\begin{aligned}
L^{\mathrm{SP}}(g_0, \alpha) = \int_0^\infty e^{-\rho t} \Bigg\{ \iint \Bigg\{ & \left[u(c + \alpha h_c) - \mu(c + \alpha h_c) - v(N + \alpha h_N) - \frac{\delta}{2}(\pi^w + \alpha h_\pi)^2 \right] (g + \alpha h_g) \\
& - (V + \alpha h_V)\phi_t + (V + \alpha h_V)\mathcal{A}^*(\alpha)\phi \\
& + \phi \left[u(c + \alpha h_c) - v(N + \alpha h_N) - \frac{\delta}{2}(\pi^w + \alpha h_\pi)^2 \right] \\
& + \chi u'(c + \alpha h_c) + \chi_a (V + \alpha h_V) \\
& + (g + \alpha h_g)\lambda_t - \rho\lambda(g + \alpha h_g) + (g + \alpha h_g)\mathcal{A}(\alpha)\lambda \Bigg\} da dz \\
& + \mu A(N + \alpha h_N) + \vartheta_t(\pi^w + \alpha h_\pi) \\
& + \vartheta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon}(1 + \tau^L) A \iint u'(c + \alpha h_c)(g + \alpha h_g) da dz - v'(N + \alpha h_N) \right) (N + \alpha h_N) \Bigg\} dt
\end{aligned}
$$

We now differentiate and take the limit $\alpha \to 0$. Setting the resulting expression to 0, we have the following first-order necessary condition for optimality:

$$
\begin{aligned}
0 = \int_0^\infty e^{-\rho t} \Bigg\{ \iint \Bigg\{ & \left[u'(c)h_c - \mu h_c - v'(N)h_N - \delta\pi^w h_\pi \right] g + h_g \left[u(c) - \mu c - v(N) - \frac{\delta}{2}(\pi^w)^2 \right] \\
& - h_V\phi_t + h_V\mathcal{A}^*(0)\phi + V\frac{d}{d\alpha}\mathcal{A}^*(0)\phi + \phi\left[u'(c)h_c - v'(N)h_N - \delta\pi^w h_\pi \right] \\
& + \chi u''(c)h_c + \chi_a h_V \\
& + h_g\lambda_t - \rho\lambda h_g + h_g\mathcal{A}(0)\lambda + g\frac{d}{d\alpha}\mathcal{A}(0)\lambda \Bigg\} da dz \\
& + \mu A h_N + \vartheta_t h_\pi \\
& + \vartheta\frac{\epsilon}{\delta}\left(\frac{\epsilon - 1}{\epsilon}(1 + \tau^L) A \iint [u'(c)h_g + u''(c) g h_c] da dz - v'(N)h_N \right) N \\
& + h_N \vartheta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon}(1 + \tau^L) A \Lambda - v'(N) \right) \Bigg\} dt
\end{aligned}
$$

where we have

$$
\frac{d}{d\alpha}\mathcal{A}(0) = (a h_r + z h_w N + z w h_N - h_c)\partial_a
$$

and, again dropping boundary terms,

$$V \frac{d}{d\alpha} \mathcal{A}^*(0)\phi = \phi \frac{d}{d\alpha} \mathcal{A}(0) V$$
$$= \phi (a h_r + z h_w N + z w h_N - h_c) V_a.$$

Finally, we group terms by h_c, h_g, etc., and invoke the fundamental lemma of the calculus of variations. We directly obtain the optimality conditions that characterize the optimal Ramsey plan of Proposition 4 in the interior of the state space.

In the following, we provide a formal treatment of boundary conditions. To that end, we start by discretizing the conditions for competitive equilibrium in the spirit of Achdou et al. (2021).

A.3 Discretized Competitive Equilibrium Conditions

We now prove a discretized representation of competitive equilibria in our baseline HANK model. This will elucidate how boundary conditions are treated formally by the Ramsey planner. For any function $c_t(a,z)$, we discretize both in the state space and in time, so that we write $\boldsymbol{c}_n$ for $n = 0, \ldots, N$. In particular, $\boldsymbol{c}_n$ is a $J \times 1$ vector, so that $c_{i,n} = c_{t_n}(a_i, z_i)$ associated with grid point i. We also use notation $\boldsymbol{c}_{n,[2:J]}$, for example, to denote the $(J-1) \times 1$ vector consisting of elements 2 through J in $\boldsymbol{c}_n$.

We summarize the discretized competitive equilibrium conditions of our model in the following Lemma, using a finite-difference discretization given a policy path $\boldsymbol{\theta} = \{\theta_n\}_{n \geq 0}$. The proof follows along the lines of Achdou et al. (2021) and Schaab and Zhang (2021), and we refer the interested reader to those papers. This characterization will justify setting up the Ramsey problem using the following discretized equations as implementability conditions.

Lemma 14. *A consistent finite-difference discretization of the implementability conditions of our baseline HANK model is as follows. For the Hamilton-Jacobi-Bellman equation, we have*

$$\rho \boldsymbol{V}_n = \frac{\boldsymbol{V}_{n+1} - \boldsymbol{V}_n}{dt} + u \begin{pmatrix} i_n a_1 - \pi_n^w a_1 + \frac{A_{n+1} - A_n}{dt A_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} - v(N_n) - \frac{\delta}{2} (\pi_n^w)^2$$
$$+ \begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi_n^w \boldsymbol{a}_{[2:J]} + \frac{A_{n+1} - A_n}{dt A_n} \boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \cdot \frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_n + \boldsymbol{A}^z \boldsymbol{V}_n$$

For the consumption first-order condition of the household, we simply have

$$u'(\boldsymbol{c}_{n,[2:J]}) = \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{n+1} \right)_{[2:J]}$$

For the Kolmogorov forward equation, we have

$$\frac{\boldsymbol{g}_{n+1}-\boldsymbol{g}_n}{dt} =(\boldsymbol{A}^z)'\boldsymbol{g}_n + \frac{\boldsymbol{D}_a'}{da}\left[\begin{pmatrix} 0 \\ i_n\boldsymbol{a}_{[2:J]} - \pi_n^w\boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n}\boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]}A_nN_n - \boldsymbol{c}_{n,[2:J]}\end{pmatrix}\cdot \boldsymbol{g}_n\right]$$

Finally, for the resource constraint we simply have

$$A_nN_n = \boldsymbol{c}_n'\boldsymbol{g}_n d\boldsymbol{x}$$

and for the Phillips curve

$$\frac{\pi_{n+1}^w - \pi_n^w}{dt} = \rho\pi_n^w + \frac{\epsilon}{\delta}\left[\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n(\boldsymbol{z}\cdot u'(\boldsymbol{c}_n))'\boldsymbol{g}_n d\boldsymbol{x} - v'(N_n)\right]N_n$$

and we have already used $c_{n,1} = i_na_1 - \pi_n^w a_1 + \frac{A_{n+1}-A_n}{dtA_n}a_1 + z_1A_nN_n$.

Crucially, the discretized system of equations in the above Lemma properly accounts for the household borrowing constraint, leveraging results from Achdou et al. (2021). In particular, they prove that in the simple Huggett economy with two earnings states the only point in the state space where the borrowing constraint binds is $(\underline{a}, z^L)$. We use this result here to plug in the borrowing constraint directly at that discretized point. While we have not formally proven that their representation extends to our HANK economy, we verify its validity numerically ex-post. And since the stationary equilibrium of our model is almost identical to theirs, there is little reason to expect any sharp discrepancies in this context.

A.4 Discretized Standard Ramsey Problem

The standard primal Ramsey problem in our baseline HANK model is associated with the discretized Lagrangian

$$
\begin{aligned}
L^{\text{SP}}(\boldsymbol{g}_0) = & \min_{\{\boldsymbol{\phi}_n,\boldsymbol{\chi}_n,\boldsymbol{\lambda}_n,\mu_n,\vartheta_n\}} \max_{\{\boldsymbol{V}_n,\boldsymbol{c}_{n,[2:J]},\boldsymbol{g}_n,\pi^w_n,N_n,i_n\}} \sum_{n=0}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ \\
& + u \begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_t - v(N_n)\mathbf{1}'\boldsymbol{g}_t - \frac{\delta}{2}(\pi^w_n)^2 \mathbf{1}'\boldsymbol{g}_t \\
& + \boldsymbol{\nu}'_n \Bigg[-\rho \boldsymbol{V}_n + \frac{\boldsymbol{V}_{n+1} - \boldsymbol{V}_n}{dt} + u \begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} - v(N_n) - \frac{\delta}{2}(\pi^w_n)^2 \Bigg] \\
& + \boldsymbol{\nu}'_n \boldsymbol{A}^z \boldsymbol{V}_n + \sum_{i \geq 2} \nu_{i,n} \left(i_n a_i - \pi^w_n a_i + \frac{A_{n+1} - A_n}{dtA_n} a_i + z_i A_n N_n - c_{n,i} \right) \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n \\
& + \boldsymbol{\chi}'_{n,[2:J]} \left[u'(\boldsymbol{c}_{n,[2:J]}) - \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{n+1} \right)_{[2:J]} \right] \\
& - \boldsymbol{\omega}'_n \frac{\boldsymbol{g}_{n+1} - \boldsymbol{g}_n}{dt} + \boldsymbol{\omega}'_n (\boldsymbol{A}^z)' \boldsymbol{g}_n \\
& + \sum_{i \geq 2} \omega_{n,i} \frac{\boldsymbol{D}'_{a,[i,:]}}{da} \left[\begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi^w_n \boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n} \boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \cdot \boldsymbol{g}_n \right] \Bigg\} d\boldsymbol{x} \\
& + \mu_n \left[\boldsymbol{c}'_n \boldsymbol{g}_n d\boldsymbol{x} - A_n N_n \right] \\
& + \vartheta_n \left[- \frac{\pi^w_{n+1} - \pi^w_n}{dt} + \rho \pi^w_n + \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A_n (\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) N_n \right] \Bigg\} dt
\end{aligned}
$$

where the planner takes as given an initial condition for the cross-sectional distribution, $\boldsymbol{g}_0$.

As in Appendix A.3, we fix from the beginning that unemployed households at the borrowing constraint always consume their income, that is

$$
c_{n,1} = i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1} - A_n}{dtA_n} a_1 + z_1 A_n N_n
$$

for all n. The planner takes this as given and does not get to consider perturbations in $c_{n,1}$ for any n.

We want to emphasize at this point how important it is exactly which finite-difference stencils are used for the discretization. For discretization in the time dimension, for example, the above Lagrangian assumes a *semi-implicit backwards* discretization of $\partial_t \boldsymbol{V}_t$ in the HJB. And it assumes an *explicit forwards* discretization of $\partial_t \boldsymbol{g}_t$ in the KF equation. For the aggregates, it assumes an *explicit forwards* discretization for $\dot{A}_t$ and also an *explicit forwards* discretization for $\dot{\pi}^w_t$. These assumptions also correspond to the appropriate stencils we use numerically to implement our results.

We also want to echo Achdou et al. (2021) at this point, recalling that the correct discretization stencil for the KF equation in the wealth dimension is given by

$$(\boldsymbol{A}^a)'\boldsymbol{g} = \frac{1}{da}(\boldsymbol{s}\cdot\boldsymbol{D}_a)'\boldsymbol{g} = \frac{1}{da}\boldsymbol{D}_a'(\boldsymbol{s}\cdot\boldsymbol{g}).$$

That is, the correct stencil uses the tranpose $\boldsymbol{D}_a'$ rather than, as one might have expected, $-\boldsymbol{D}_a(\boldsymbol{s}\cdot\boldsymbol{g})$.

A.5 Auxilliary Results

Before tackling the main proof of this appendix, we state several auxilliary results that will be helpful below. Most of these results follow trivially by applying well-known properties of matrix algebra. We consequently provide only some of the proofs explicitly.

Lemma 15. *The following matrix algebra tricks will be useful. Let $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ be $J\times 1$ vectors and $\boldsymbol{A}$ a $J\times J$ matrix. Transposition satisfies*

$$(\boldsymbol{A}\boldsymbol{x})' = \boldsymbol{x}'\boldsymbol{A}'.$$

We also have

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y} = \sum_i x_i \boldsymbol{A}_{[i,:]}\boldsymbol{y} = \sum_i x_i \sum_j \boldsymbol{A}_{[i,j]} y_j = \sum_j y_j \sum_i \boldsymbol{A}'_{[j,i]} x_i = \boldsymbol{y}'\boldsymbol{A}'\boldsymbol{x}.$$

We also have

$$\boldsymbol{x}'(\boldsymbol{y}\cdot\boldsymbol{A})\boldsymbol{z} = \boldsymbol{x}'(\boldsymbol{y}\cdot(\boldsymbol{A}\boldsymbol{z})) = (\boldsymbol{x}\cdot\boldsymbol{y})'\boldsymbol{A}z = (\boldsymbol{A}\boldsymbol{z})'(\boldsymbol{x}\cdot\boldsymbol{y}) = \boldsymbol{z}'\boldsymbol{A}'(\boldsymbol{x}\cdot\boldsymbol{y}) = \boldsymbol{z}'(\boldsymbol{y}\cdot\boldsymbol{A})'\boldsymbol{x}.$$

Taking derivatives, we have

$$\frac{d}{d\boldsymbol{x}}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}\boldsymbol{y}$$
$$\frac{d}{d\boldsymbol{x}}\boldsymbol{y}'\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{y}'\boldsymbol{A})' = \boldsymbol{A}'\boldsymbol{y}$$

Lemma 16. *In the Lagrangian, the HJB term can be rearranged as follows:*

$$
\begin{aligned}
\frac{1}{da}\sum_{i\geq 2}\nu_i s_i \boldsymbol{D}_{a,[i,:]}\boldsymbol{V} &= \frac{1}{da}\sum_{i\geq 1}\nu_i s_i \boldsymbol{D}_{a,[i,:]}\boldsymbol{V} \\
&= \frac{1}{da}\boldsymbol{\nu}'(\boldsymbol{s}\cdot\boldsymbol{D}_a)\boldsymbol{V} \\
&= \frac{1}{da}\boldsymbol{V}'(\boldsymbol{s}\cdot\boldsymbol{D}_a)'\boldsymbol{\nu} \\
&= \frac{1}{da}\boldsymbol{V}'\boldsymbol{D}_a'(\boldsymbol{s}\cdot\boldsymbol{\nu})
\end{aligned}
$$

where $\boldsymbol{D}_a$ is the upwind finite-difference matrix in the a dimension. We sometimes use $\boldsymbol{s}\cdot\boldsymbol{D}_a = \boldsymbol{A}^a$.

Proof. We have

$$
\begin{aligned}
\sum_{i\geq 2}\nu_i s_i \boldsymbol{D}_{a,[i,\,:]}\boldsymbol{V} &= \sum_{i\geq 2}\nu_i s_i \sum_{j\geq 1}\boldsymbol{D}_{a,[i,\,j]}\boldsymbol{V}_j \\
&= \sum_{i\geq 2}\nu_i s_i \sum_{j\geq 1}\boldsymbol{D}'_{a,[j,\,i]}\boldsymbol{V}_j \\
&= \sum_{j\geq 1}\boldsymbol{V}_j \sum_{i\geq 2}\boldsymbol{D}'_{a,[j,\,i]}\nu_i s_i \\
&= \sum_{j\geq 1}\boldsymbol{V}_j \sum_{i\geq 1}\boldsymbol{D}'_{a,[j,\,i]}\nu_i s_i \\
&= \sum_{j\geq 1}\boldsymbol{V}_j \boldsymbol{D}'_{a,[j,\,:]}(\boldsymbol{s}\cdot\boldsymbol{\nu}) \\
&= \boldsymbol{V}'\boldsymbol{D}'_a(\boldsymbol{s}\cdot\boldsymbol{\nu})
\end{aligned}
$$

where $D'_{a,[j,\,:]}$ denotes the jth row of the matrix $\boldsymbol{D}'_a$. ■

Lemma 17. *The correct adjoint operation, i.e., the one we use to define $\mathcal{A}^* \approx \boldsymbol{A}'$, is given by*

$$\boldsymbol{D}'_a(\boldsymbol{s}\cdot\boldsymbol{\nu}) = (\boldsymbol{A}^a)'\boldsymbol{\nu}.$$

In particular, we have

$$\begin{aligned}
\boldsymbol{\omega}'(\boldsymbol{A}^a)'\boldsymbol{g} &= \boldsymbol{\omega}'\boldsymbol{D}_a'(\boldsymbol{s}\cdot\boldsymbol{g}) \\
&= (\boldsymbol{s}\cdot\boldsymbol{g})'\boldsymbol{D}_a\boldsymbol{\omega} \\
&= \boldsymbol{g}'(\boldsymbol{s}\cdot\boldsymbol{D}_a)\boldsymbol{\omega} \\
&= (\boldsymbol{D}_a\boldsymbol{\omega})'(\boldsymbol{s}\cdot\boldsymbol{g}) \\
&= (\boldsymbol{D}_a\boldsymbol{\omega})'\left[\begin{pmatrix} 0 \\ r\boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} - \boldsymbol{c}_{[2:J]} - G \end{pmatrix}\cdot\boldsymbol{g}\right].
\end{aligned}$$

Lemma 18. *We can "integrate by parts" the FOC term in the Lagrangian to arrive at*

$$\frac{1}{da}\boldsymbol{\chi}'_{t(n),[2:J]}\Big(\boldsymbol{D}_a\boldsymbol{V}_{t(n+1)}\Big)_{[2:J]} = \frac{1}{da}\boldsymbol{V}'_{t(n+1)}\boldsymbol{D}'_a\begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J],t(n)} \end{pmatrix}.$$

Proof. We have

$$\begin{aligned}
\frac{1}{da}\sum_{i\geq 2}\chi_{i,t(n)}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}_{t(n+1)} &= \frac{1}{da}\sum_{i\geq 2}\chi_{i,t(n)}\sum_{j\geq 1}\boldsymbol{D}_{a,[i,j]}\boldsymbol{V}_{j,t(n+1)} \\
&= \frac{1}{da}\sum_{j\geq 1}\boldsymbol{V}_{j,t(n+1)}\sum_{i\geq 2}\boldsymbol{D}'_{a,[j,i]}\chi_{i,t(n)} \\
&= \frac{1}{da}\sum_{j\geq 1}\boldsymbol{V}_{j,t(n+1)}\boldsymbol{D}'_{a,[j,:]}\begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J],t(n)} \end{pmatrix} \\
&= \frac{1}{da}\boldsymbol{V}'_{t(n+1)}\boldsymbol{D}'_a\begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J],t(n)} \end{pmatrix}.
\end{aligned}$$

It is important to note that we *cannot* roll the sum $\sum_{i\geq 2}$ forward to simply read $\sum_{i\geq 1}$. This is only possible for the terms that include savings, using the fact that $s_1 = 0$. ■

Lemma 19. *We can "integrate by parts" in the time dimension as follows. For any $\boldsymbol{x}_n$, we have*

$$\sum_{n=0}^{N-1} e^{-\rho t_n}\boldsymbol{x}_{n+1} = e^{\rho dt}\sum_{n=0}^{N-1} e^{-\rho t_n}\boldsymbol{x}_n - e^{\rho dt}\boldsymbol{x}_0 + e^{\rho dt}e^{-\rho t_N}\boldsymbol{x}_N.$$

We prove the following results below. In particular, this implies

$$\sum_{n=0}^{N-1} e^{-\rho t_n} \boldsymbol{\nu}_n' \boldsymbol{V}_{n+1} = \sum_{n=0}^{N-1} e^{-\rho t_n} e^{\rho dt} \boldsymbol{\nu}_{n-1}' \boldsymbol{V}_n - e^{\rho dt} \boldsymbol{\nu}_{-1}' \boldsymbol{V}_0 + e^{\rho dt} e^{-\rho t_N} \boldsymbol{\nu}_{N-1} \boldsymbol{V}_N$$

as well as

$$-\sum_{n=0}^{N-1} e^{-\rho t_n} \boldsymbol{\omega}_n' \frac{\boldsymbol{g}_{n+1} - \boldsymbol{g}_n}{dt} = \sum_{n=0}^{N-1} e^{-\rho t_n} \frac{\boldsymbol{\omega}_n' - e^{\rho dt} \boldsymbol{\omega}_{n-1}'}{dt} \boldsymbol{g}_n + \frac{1}{dt} e^{\rho dt} \boldsymbol{\omega}_{-1}' \boldsymbol{g}_0 - \frac{1}{dt} e^{\rho dt} e^{-\rho t_N} \boldsymbol{\omega}_{N-1}' \boldsymbol{g}_N$$

and

$$\sum_{n=0}^{N-1} e^{-\rho t_n} \boldsymbol{\chi}_{n,[2:J]}' \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{n+1} \right)_{[2:J]} = \sum_{n=0}^{N-1} e^{-\rho t_n} e^{\rho dt} \boldsymbol{\chi}_{n-1,[2:J]}' \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_n \right)_{[2:J]} - e^{\rho dt} \boldsymbol{\chi}_{-1,[2:J]}' \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_0 \right)_{[2:J]} + e^{\rho dt} e^{-\rho t_N} \boldsymbol{\chi}_{N-1,[2:J]}' \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_N \right)_{[2:J]}$$

Finally, we have

$$-\sum_{n=0}^{N-1} e^{-\rho t_n} \vartheta_n \frac{\pi_{n+1}^w - \pi_n^w}{dt} = \sum_{n=0}^{N-1} e^{-\rho t_n} \frac{\vartheta_n - e^{\rho dt} \vartheta_{n-1}}{dt} \pi_n^w + \frac{1}{dt} e^{\rho dt} \vartheta_{-1} \pi_0^w - \frac{1}{dt} e^{\rho dt} e^{-\rho t_N} \vartheta_{N-1} \pi_N^w$$

Proof. We have

$$\begin{aligned}
\sum_{n=0}^{\infty} e^{-\rho t(n)} \boldsymbol{\nu}_{t(n)}' \frac{1}{dt} \boldsymbol{V}_{t(n+1)} &= \sum_{n=0}^{\infty} e^{-\rho t(n)} e^{\rho t(n+1)} e^{-\rho t(n+1)} \boldsymbol{\nu}_{t(n)}' \frac{1}{dt} \boldsymbol{V}_{t(n+1)} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n+1)} e^{\rho t(n+1) - \rho t(n)} \boldsymbol{\nu}_{t(n)}' \frac{1}{dt} \boldsymbol{V}_{t(n+1)} \\
&= \sum_{n=1}^{\infty} e^{-\rho t(n)} e^{\rho dt} \boldsymbol{\nu}_{t(n-1)}' \frac{1}{dt} \boldsymbol{V}_{t(n)} \\
&= \sum_{n=1}^{\infty} e^{-\rho t(n)} e^{\rho dt} \boldsymbol{\nu}_{t(n-1)}' \frac{1}{dt} \boldsymbol{V}_{t(n)} + e^{-\rho t(0)} e^{\rho dt} \boldsymbol{\nu}_{-1}' \frac{1}{dt} \boldsymbol{V}_{t(0)} - e^{-\rho t(0)} e^{\rho dt} \boldsymbol{\nu}_{-1}' \frac{1}{dt} \boldsymbol{V}_{t(0)} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n)} e^{\rho dt} \boldsymbol{\nu}_{t(n-1)}' \frac{1}{dt} \boldsymbol{V}_{t(n)} - e^{-\rho t(0)} e^{\rho dt} \boldsymbol{\nu}_{-1}' \frac{1}{dt} \boldsymbol{V}_{t(0)}.
\end{aligned}$$

Similarly, we can rearrange

$$
\begin{aligned}
\sum_{n=0}^{\infty} e^{-\rho t(n)} \boldsymbol{\chi}'_{t(n),[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(n+1)} \right)_{[2:J]} &= \sum_{n=0}^{\infty} e^{-\rho t(n)} e^{\rho t(n+1)} e^{-\rho t(n+1)} \boldsymbol{\chi}'_{t(n),[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(n+1)} \right)_{[2:J]} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n+1)} e^{\rho t(n+1) - \rho t(n)} \boldsymbol{\chi}'_{t(n),[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(n+1)} \right)_{[2:J]} \\
&= \sum_{n=1}^{\infty} e^{-\rho t(n)} e^{\rho dt} \boldsymbol{\chi}'_{t(n-1),[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(n)} \right)_{[2:J]} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n)} e^{\rho dt} \boldsymbol{\chi}'_{t(n-1),[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(n)} \right)_{[2:J]} - e^{-\rho t(0)} e^{\rho dt} \boldsymbol{\chi}'_{-1,[2:J]} \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{t(0)} \right)_{[2:J]}
\end{aligned}
$$

Finally, notice that

$$
\begin{aligned}
e^{\rho dt} \boldsymbol{\nu}'_{t(n-1)} \frac{1}{dt} \boldsymbol{V}_{t(n)} &= (1 + \rho dt) \boldsymbol{\nu}'_{t(n-1)} \frac{1}{dt} \boldsymbol{V}_{t(n)} \\
&= \boldsymbol{\nu}'_{t(n-1)} \frac{1}{dt} \boldsymbol{V}_{t(n)} + \rho \boldsymbol{\nu}'_{t(n-1)} \boldsymbol{V}_{t(n)}
\end{aligned}
$$

Lastly,

$$
\begin{aligned}
-\sum_{n=0}^{\infty} e^{-\rho t(n)} \boldsymbol{\omega}'_{t(n)} \frac{\boldsymbol{g}_{t(n+1)} - \boldsymbol{g}_{t(n)}}{dt(n)} &= -\sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n+1)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)} \\
&= -\frac{1}{dt} \sum_{n=0}^{\infty} e^{-\rho t(n)} e^{\rho t(n+1)} e^{-\rho t(n+1)} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n+1)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)} \\
&= -\frac{1}{dt} e^{\rho dt} \sum_{n=0}^{\infty} e^{-\rho t(n+1)} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n+1)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)} \\
&= -\frac{1}{dt} e^{\rho dt} \sum_{n=1}^{\infty} e^{-\rho t(n)} \boldsymbol{\omega}'_{t(n-1)} \boldsymbol{g}_{t(n)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)}
\end{aligned}
$$

And so we get

$$
\begin{aligned}
&= -\frac{1}{dt} e^{\rho dt} \sum_{n=1}^{\infty} e^{-\rho t(n)} \boldsymbol{\omega}'_{t(n-1)} \boldsymbol{g}_{t(n)} + \frac{1}{dt} e^{\rho dt} e^{-\rho t(0)} \boldsymbol{\omega}'_{t(-1)} \boldsymbol{g}_{t(0)} - \frac{1}{dt} e^{\rho dt} e^{-\rho t(0)} \boldsymbol{\omega}'_{t(-1)} \boldsymbol{g}_{t(0)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)} \\
&= -\frac{1}{dt} e^{\rho dt} \sum_{n=0}^{\infty} e^{-\rho t(n)} \boldsymbol{\omega}'_{t(n-1)} \boldsymbol{g}_{t(n)} + \frac{1}{dt} e^{\rho dt} e^{-\rho t(0)} \boldsymbol{\omega}'_{t(-1)} \boldsymbol{g}_{t(0)} + \sum_{n=0}^{\infty} e^{-\rho t(n)} \frac{1}{dt} \boldsymbol{\omega}'_{t(n)} \boldsymbol{g}_{t(n)} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n)} \left(\frac{1}{dt} \boldsymbol{\omega}'_{t(n)} - \frac{1}{dt} e^{\rho dt} \boldsymbol{\omega}'_{t(n-1)} \right) \boldsymbol{g}_{t(n)} + \frac{1}{dt} e^{\rho dt} e^{-\rho t(0)} \boldsymbol{\omega}'_{t(-1)} \boldsymbol{g}_{t(0)} \\
&= \sum_{n=0}^{\infty} e^{-\rho t(n)} \left(\frac{\boldsymbol{\omega}'_{t(n)} - \boldsymbol{\omega}'_{t(n-1)}}{dt} - \rho \boldsymbol{\omega}'_{t(n-1)} \right) \boldsymbol{g}_{t(n)} + \frac{1}{dt} e^{\rho dt} e^{-\rho t(0)} \boldsymbol{\omega}'_{t(-1)} \boldsymbol{g}_{t(0)}
\end{aligned}
$$

Finally, we drop the second term on the RHS because $\boldsymbol{g}_{t(0)}$ is fixed as an initial condition and so it does not respond to $\frac{d}{d\boldsymbol{\theta}}$, which is precisely why the KFE is not a forward-looking constraint. ■

Lemma 20. *In the continuous time limit as $dt \to 0$, we have*

$$e^{\rho dt} \approx 1 + \rho dt.$$

A.6 Proof of Proposition 4 with Boundary Condition

We are now ready to present our main proof. We use the auxilliary results above to rewrite the discretized Lagrangian that corresponds to the standard primal Ramsey problem of Section 4 as

$$
\begin{aligned}
L^{\text{SP}}(\boldsymbol{g}_0) = \min_{\{\boldsymbol{\phi}_n,\boldsymbol{\chi}_n,\lambda_n,\mu_n,\vartheta_n\}} \max_{\{\boldsymbol{V}_n,\boldsymbol{c}_{n,[2:J]},\boldsymbol{g}_n,\pi^w_n,N_n,i_n\}} \sum_{n=0}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ & u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_n \\
& + \mu_n \begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_n - v(N_n)\mathbf{1}'\boldsymbol{g}_n - \frac{\delta}{2}(\pi^w_n)^2 \mathbf{1}'\boldsymbol{g}_n \\
& - \frac{\boldsymbol{\nu}_n' - e^{\rho dt}\boldsymbol{\nu}_{n-1}'}{dt}\boldsymbol{V}_n + \boldsymbol{\nu}_n'\left[-\rho \boldsymbol{V}_n + u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} - v(N_n) - \frac{\delta}{2}(\pi^w_n)^2 \right] \\
& + \boldsymbol{\nu}_n' \boldsymbol{A}^z \boldsymbol{V}_n + \frac{1}{da}\boldsymbol{V}_n'(\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)' \begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi^w_n \boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n}\boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \\
& + \boldsymbol{\chi}_{n,[2:J]}' u'(\boldsymbol{c}_{n,[2:J]}) - e^{\rho dt}\frac{1}{da}\boldsymbol{V}_n'\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{n-1,[2:J]} \end{pmatrix} \\
& + \frac{\boldsymbol{\omega}_n' - e^{\rho dt}\boldsymbol{\omega}_{n-1}'}{dt}\boldsymbol{g}_n + \boldsymbol{\omega}_n'(\boldsymbol{A}^z)'\boldsymbol{g}_n \\
& + \frac{1}{da}(\boldsymbol{D}_a \boldsymbol{\omega}_t)' \left[\left(\begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi^w_n \boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n}\boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \cdot \boldsymbol{g}_n \right] \right\} d\boldsymbol{x} \\
& - \mu_n A_n N_n \\
& + \frac{\vartheta_n - e^{\rho dt}\vartheta_{n-1}}{dt}\pi^w_n + \vartheta_n \rho \pi^w_n - \vartheta_n \frac{\epsilon}{\delta} v'(N_n) N_n \\
& + \vartheta_n \frac{\epsilon}{\delta} N_n \frac{\epsilon - 1}{\epsilon}(1+\tau^L) A_n (\boldsymbol{z}\cdot\boldsymbol{g}_n)' u'\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} d\boldsymbol{x} \Bigg\} dt \\
& - e^{\rho dt}\boldsymbol{\nu}_{-1}'\boldsymbol{V}_0 d\boldsymbol{x} + e^{\rho dt} e^{-\rho t_N}\boldsymbol{\nu}_{N-1}\boldsymbol{V}_N d\boldsymbol{x} \\
& + e^{\rho dt}\frac{1}{da}\boldsymbol{V}_0'\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{-1,[2:J]} \end{pmatrix} d\boldsymbol{x}dt - e^{\rho dt}e^{-\rho t_N}\boldsymbol{V}_N'\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{N-1,[2:J]} \end{pmatrix} d\boldsymbol{x}dt \\
& + e^{\rho dt}\boldsymbol{\omega}_{-1}'\boldsymbol{g}_0 d\boldsymbol{x} - e^{\rho dt}e^{-\rho t_N}\boldsymbol{\omega}_{N-1}'\boldsymbol{g}_N d\boldsymbol{x} \\
& + e^{\rho dt}\vartheta_{-1}\pi^w_0 - e^{\rho dt}e^{-\rho t_N}\vartheta_{N-1}\pi^w_N
\end{aligned}
$$

In the spirit of Marcet and Marimon (2019), we have reordered the forward-looking constraints—this corresponds to integration by parts in the time dimension in the fully continuous case. The resulting "boundary" terms in the last few lines of the above Lagrangian are the key objects at the heart of the time consistency problems we discuss in Sections 4.

We are now ready to take derivatives and characterize necessary first-order conditions for the standard Ramsey plan.

Derivative V_n. We have

$$0 = -\frac{\boldsymbol{\nu}_n' - e^{\rho dt}\boldsymbol{\nu}_{n-1}'}{dt} - \rho\boldsymbol{\nu}_n + (\boldsymbol{A}^z)'\boldsymbol{\nu}_n + \frac{1}{da}(\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)'\boldsymbol{s}_n - e^{\rho dt}\frac{1}{da}\boldsymbol{D}_a'\begin{pmatrix} 0 \\ \chi_{n-1,[2:J]} \end{pmatrix}$$

Using our auxilliary results, we have $(\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)'\boldsymbol{s}_n = (\boldsymbol{s}_n \cdot \boldsymbol{D}_a)'\boldsymbol{\nu}_n = (\boldsymbol{A}^a)'\boldsymbol{\nu}_n$, and so

$$0 = -\frac{\boldsymbol{\nu}_n' - e^{\rho dt}\boldsymbol{\nu}_{n-1}'}{dt} - \rho\boldsymbol{\nu}_n + \boldsymbol{A}'\boldsymbol{\nu}_n - e^{\rho dt}\frac{1}{da}\boldsymbol{D}_a'\begin{pmatrix} 0 \\ \chi_{n-1,[2:J]} \end{pmatrix}.$$

Derivative g_n. We have

$$\begin{aligned} 0 =& u(\boldsymbol{c}_n) + \mu_n\boldsymbol{c}_n - v(N_n)\boldsymbol{1} - \frac{\delta}{2}(\pi_n^w)^2\boldsymbol{1} + \frac{\boldsymbol{\omega}_n' - e^{\rho dt}\boldsymbol{\omega}_{n-1}'}{dt} + (\boldsymbol{\omega}_n'(\boldsymbol{A}^z)')' \\ &+ \frac{d}{d\boldsymbol{g}_n}\left[\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\omega}_n)'[\boldsymbol{s}_n \cdot \boldsymbol{g}_n]\right] + \vartheta_n N_t\frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)A_n\boldsymbol{z} \cdot u'(\boldsymbol{c}_n) \end{aligned}$$

Now we work out the remaining derivative,

$$\begin{aligned} \frac{d}{d\boldsymbol{g}_n}\left[\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\omega}_n)'[\boldsymbol{s}_n \cdot \boldsymbol{g}_n]\right] &= \frac{1}{da}\frac{d}{d\boldsymbol{g}_n}\left[\left(\boldsymbol{s}_n' \cdot (\boldsymbol{D}_a\boldsymbol{\omega}_n)'\right)\boldsymbol{g}_n\right] \\ &= \frac{1}{da}\frac{d}{d\boldsymbol{g}_n}\left[\boldsymbol{g}_n'\left(\boldsymbol{s}_n \cdot (\boldsymbol{D}_a\boldsymbol{\omega}_n)\right)\right] \\ &= \frac{1}{da}\frac{d}{d\boldsymbol{g}_n}\left[\boldsymbol{g}_n'\left((\boldsymbol{s}_n \cdot \boldsymbol{D}_a)\boldsymbol{\omega}_n\right)\right] \\ &= \frac{1}{da}(\boldsymbol{s}_n \cdot \boldsymbol{D}_a)\boldsymbol{\omega}_n \\ &= \boldsymbol{A}^a\boldsymbol{\omega}_n. \end{aligned}$$

Thus, we have

$$0 = u(\boldsymbol{c}_n) + \mu_n\boldsymbol{c}_n - v(N_n)\boldsymbol{1} - \frac{\delta}{2}(\pi_n^w)^2\boldsymbol{1} + \frac{\boldsymbol{\omega}_n' - e^{\rho dt}\boldsymbol{\omega}_{n-1}'}{dt} + \boldsymbol{A}\boldsymbol{\omega}_n + \vartheta_n N_t\frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)A_n\boldsymbol{z} \cdot u'(\boldsymbol{c}_n)$$

Derivative $\boldsymbol{c}_{n,[2:J]}$. We now take the derivative with respect to $c_{n,i}$ for $i \geq 2$. We have

$$0 = u'(c_{n,i})g_{n,i} + \mu_n g_{n,i} + u'(c_{n,i})\nu_{n,i} + u''(c_{n,i})\chi_{n,i} + \vartheta_n N_n \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n z_i u''(c_{n,i})g_{n,i}$$
$$+ \frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n'(\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)\boldsymbol{V}_n\right] + \frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n' \cdot (\boldsymbol{D}_a\boldsymbol{\omega}_n)'\boldsymbol{g}_n\right]$$

Working out the remaining derivatives, we have

$$\begin{aligned}
\frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n'(\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)\boldsymbol{V}_n\right] &= \frac{1}{da}\bigg((\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)\boldsymbol{V}_n\bigg)_{[i]}\frac{ds_{n,i}}{dc_{n,i}} \\
&= -\frac{1}{da}\bigg((\boldsymbol{\nu}_n \cdot \boldsymbol{D}_a)\boldsymbol{V}_n\bigg)_{[i]} \\
&= -\frac{1}{da}\nu_{n,i}\bigg(\boldsymbol{D}_a\boldsymbol{V}_n\bigg)_{[i]} \\
&= -\frac{1}{da}\nu_{n,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}_n.
\end{aligned}$$

And similarly,

$$\begin{aligned}
\frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n' \cdot (\boldsymbol{D}_a\boldsymbol{\omega}_n)'\boldsymbol{g}_n\right] &= \frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{g}_n'\bigg(\boldsymbol{s}_n \cdot (\boldsymbol{D}_a\boldsymbol{\omega}_n)\bigg)\right] \\
&= \frac{ds_{n,it}}{dc_{n,i}}\frac{1}{da}g_{n,i}(\boldsymbol{D}_a\boldsymbol{\omega}_n)_{[i]} \\
&= -\frac{1}{da}g_{n,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{\omega}_n.
\end{aligned}$$

Thus, we have

$$0 = u'(c_{n,i})g_{n,i} + \mu_n g_{n,i} + u'(c_{n,i})\nu_{n,i} + u''(c_{n,i})\chi_{n,i} + \vartheta_n N_n \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n z_i u''(c_{n,i})g_{n,i}$$
$$- \frac{1}{da}\nu_{n,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}_n - \frac{1}{da}g_{n,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{\omega}_n.$$

Derivative π_n^w. We have

$$0 = \Big[-u'(c_{n,1})g_{n,1}a_1 - \mu_n g_{n,1}a_1 - \delta\pi_n^w \mathbf{1}'\boldsymbol{g}_n - \nu_{n,1}u'(c_{n,1})a_1 - \delta\pi_n^w \boldsymbol{\nu}_n'\mathbf{1}\Big] d\boldsymbol{x}$$

$$- \vartheta_n \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n N_n z_1 u''(c_{n,1})g_{n,1}a_1 d\boldsymbol{x}$$

$$+ \Big[-\sum_{i\geq 2}\nu_{i,n}a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_n + \sum_{i\geq 2}\omega_{n,i}\frac{\boldsymbol{D}'_{a,[i,:]}}{da}\Big[\begin{pmatrix}0\\ -\boldsymbol{a}_{[2:J]}\end{pmatrix}\cdot \boldsymbol{g}_n\Big] d\boldsymbol{x}$$

$$+ \frac{\vartheta_n - e^{\rho dt}\vartheta_{n-1}}{dt} + \rho\vartheta_n$$

Alternatively, we have

$$\begin{aligned}
\frac{d}{d\pi_n^w}\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\omega}_n)'[\boldsymbol{s}_n\cdot\boldsymbol{g}_n] &= \frac{d}{d\pi_n^w}\frac{1}{da}\Big(\boldsymbol{s}_n\cdot\boldsymbol{D}_a\boldsymbol{\omega}_n\Big)'\boldsymbol{g}_n \\
&= \frac{d}{d\pi_n^w}\frac{1}{da}\boldsymbol{g}_n'\Big(\boldsymbol{s}_n\cdot\boldsymbol{D}_a\boldsymbol{\omega}_n\Big) \\
&= \frac{1}{da}\boldsymbol{g}_n'\Big(\frac{d\boldsymbol{s}_n}{d\pi_n^w}\cdot\boldsymbol{D}_a\boldsymbol{\omega}_n\Big) \\
&= \frac{1}{da}\boldsymbol{g}_n'\left(\begin{pmatrix}0\\ -\boldsymbol{a}_{[2:J]}\end{pmatrix}\cdot\boldsymbol{D}_a\boldsymbol{\omega}_n\right) \\
&= \sum_{i\geq 1} g_{n,i}\left(\begin{pmatrix}0\\ -\boldsymbol{a}_{[2:J]}\end{pmatrix}\cdot\frac{\boldsymbol{D}_a}{da}\boldsymbol{\omega}_n\right)_{[i]} \\
&= -\sum_{i\geq 2} g_{n,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}_n
\end{aligned}$$

Thus, we have

$$0 = \Big[-u'(c_{n,1})g_{n,1}a_1 - \mu_n g_{n,1}a_1 - \delta\pi_n^w \mathbf{1}'\boldsymbol{g}_n - \nu_{n,1}u'(c_{n,1})a_1 - \delta\pi_n^w \boldsymbol{\nu}_n'\mathbf{1}\Big] d\boldsymbol{x}$$

$$- \vartheta_n \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n N_n z_1 u''(c_{n,1})g_{n,1}a_1 d\boldsymbol{x}$$

$$-\sum_{i\geq 2}\nu_{i,n}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_n d\boldsymbol{x} - \sum_{i\geq 2}g_{n,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}_n d\boldsymbol{x} + \frac{\vartheta_n - e^{\rho dt}\vartheta_{n-1}}{dt} + \rho\vartheta_n$$

Derivative i_n. The nominal interest rate derivative is very easy because it's parallel to wage inflation, except in the Phillips curve. That is, we have

$$
\begin{aligned}
0 =& u'(c_{n,1})g_{n,1}a_1 + \mu_n g_{n,1}a_1 + \nu_{n,1}u'(c_{n,1})a_1 + \sum_{i\geq 2}\nu_{i,n}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_n + \sum_{i\geq 2}g_{n,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}_n \\
&+ \vartheta_n\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_nN_nz_1u''(c_{n,1})g_{n,1}a_1
\end{aligned}
$$

Derivative N_n. Finally, we take the derivative for aggregate labor. This yields

$$
\begin{aligned}
0 =& \left[u'(c_{n,1})g_{n,1}z_1A_n + \mu_n g_{n,1}z_1A_n + \nu_{n,1}u'(c_{n,1})z_1A_n + \sum_{i\geq 2}\nu_{i,n}z_iA_n\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_n + \sum_{i\geq 2}g_{n,i}z_iA_n\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}_n\right]d\boldsymbol{x} \\
&+ \vartheta_n\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_nN_nz_1u''(c_{n,1})g_{n,1}z_1A_nd\boldsymbol{x} \\
&- v'(N_n)\boldsymbol{1}'\boldsymbol{g}_nd\boldsymbol{x} - v'(N_n)\boldsymbol{\nu}_n'\boldsymbol{1}d\boldsymbol{x} \\
&- \mu_nA_n + \vartheta_n\frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n(\boldsymbol{z}\cdot u'(\boldsymbol{c}_n))'\boldsymbol{g}_nd\boldsymbol{x} - v'(N_n)\right) - \frac{\epsilon}{\delta}\vartheta_nv''(N_n)N_n
\end{aligned}
$$

These derivations conclude our proof. In particular, the first-order conditions we have now derived are the exact, discretized analogs of the conditions we present in Proposition 4. For convenience, we formally state the discretized characterization of the stationary Ramsey plan here, so that interested readers can follow the mapping more easily. In the following representation of the stationary Ramsey plan, we use the fact that, in any stationary equilibrium, we simply have

$$
u'(\boldsymbol{c}_i) = \frac{1}{da}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}
$$

for $i \geq 2$.

Lemma 21. *(Discretized Stationary Ramsey Plan) A consistent discretization of the stationary Ramsey plan, with $A_{ss} = 1$, is given by the following equations. For the value function, we have*

$$
0 = -\frac{1-e^{\rho dt}}{dt}\boldsymbol{\nu} - \rho\boldsymbol{\nu} + \boldsymbol{A}'\boldsymbol{\nu} - e^{\rho dt}\frac{1}{da}\boldsymbol{D}_a'\begin{pmatrix}0\\ \boldsymbol{\chi}_{[2:J]}\end{pmatrix}
$$

and for the distribution

$$
0 = \frac{1-e^{\rho dt}}{dt}\boldsymbol{\omega} + \boldsymbol{A}\boldsymbol{\omega} + u(\boldsymbol{c}) + \mu\boldsymbol{c} - v(N) - \frac{\delta}{2}(\pi^w)^2 + \vartheta N\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)\boldsymbol{z}\cdot u'(\boldsymbol{c})
$$

For consumption, for $i \geq 2$, we have

$$-u''(c_i)\chi_i = \left[u'(c_i) + \mu + \vartheta N \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L) z_i u''(c_i) - \frac{1}{da}\boldsymbol{D}_{a,[i,:]}\boldsymbol{\omega}\right] g_i$$

The optimality condition for monetary policy, i.e., the nominal interest rate, is given by

$$0 = \left(u'(c_1) + \mu - \frac{1}{da}\boldsymbol{D}_{a,[1,:]}\boldsymbol{\omega} + \vartheta \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L) N z_1 u''(c_1)\right) g_1 a_1 + \sum_{i \geq 1} \nu_i a_i u'(c_i) + \sum_{i \geq 1} g_i a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}$$

We see here nicely how we need a boundary correction at the borrowing constraint. For inflation, we have

$$0 = -\delta\pi^w - \delta\pi^w \boldsymbol{\nu}'\mathbf{1} d\boldsymbol{x} + \frac{1 - e^{\rho dt}}{dt}\vartheta + \rho\vartheta$$

where we used the optimality condition for monetary policy to drop terms. Finally, the optimality condition for aggregate labor, i.e., aggregate economic activity, is given by

$$\begin{aligned} 0 = &\left[\left(u'(c_1) + \mu - \frac{1}{da}\boldsymbol{D}_{a,[1,:]}\boldsymbol{\omega} + \vartheta \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L) N z_1 u''(c_1)\right) g_1 z_1 + \sum_{i \geq 1} \nu_i z_i u'(c_i) + \sum_{i \geq 1} g_i z_i \frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\omega}\right] d\boldsymbol{x} \\ &- v'(N) - v'(N)\boldsymbol{\nu}'\mathbf{1} d\boldsymbol{x} - \mu + \vartheta\frac{\epsilon}{\delta}\left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L)(\boldsymbol{z} \cdot u'(\boldsymbol{c}))'\boldsymbol{g} d\boldsymbol{x} - v'(N)\right) - \frac{\epsilon}{\delta}\vartheta v''(N) N \end{aligned}$$

Remark 22. Crucially, this discretized representation of the Ramsey plan provides a formal treatment of boundary conditions. We see exactly how the planner takes into account the borrowing constraint that households face. And we see exactly where the corresponding boundary terms enter the optimality conditions and targeting rules for optimal monetary policy.

A.7 Proof of Proposition 6

Our goal is to show that

$$\frac{dL^{\text{TD}}(\boldsymbol{g}_{\text{ss}}, \boldsymbol{\phi}_{\text{ss}}, \vartheta_{\text{ss}}, \boldsymbol{\theta}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}})}{d\boldsymbol{\theta}} = F\left(\boldsymbol{g}_{\text{ss}}, \boldsymbol{\phi}_{\text{ss}}, \vartheta_{\text{ss}}, \boldsymbol{\theta}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}\right) = 0. \tag{55}$$

In the following, we will prove that this perturbation is 0 for a given $\frac{d}{d\theta_k}$, and we can then simply "stack" up to arrive at any perturbation $\frac{d}{d\boldsymbol{\theta}}$. For our baseline HANK model, $\frac{dL^{\text{TD}}}{d\theta_k}$ takes the form

$$0 = \frac{d}{d\theta_k}\Bigg\{ \sum_{n=0}^{\infty} e^{-\rho t}\Bigg\{ u\begin{pmatrix} i_n a_1 - \pi_n^w a_1 + \frac{A_{n+1}-A_n}{dtA_n}a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_n - v(N_n)\mathbf{1}'\boldsymbol{g}_n - \frac{\delta}{2}(\pi_n^w)^2\mathbf{1}'\boldsymbol{g}_n \Bigg\} d\boldsymbol{x}dt$$

$$\underbrace{+\frac{1}{dt}e^{\rho dt}\boldsymbol{\phi}'\boldsymbol{V}_0 d\boldsymbol{x} - \frac{1}{dt}e^{\rho dt}\vartheta\pi_0^w}_{\text{Timeless Penalties}}\Bigg\}\Bigg|_{\boldsymbol{g}_{\text{ss}},\boldsymbol{\phi}_{\text{ss}},\vartheta_{\text{ss}},\boldsymbol{\theta}_{\text{ss}},\boldsymbol{Z}_{\text{ss}}}$$

for all $k \geq 0$. We start by evaluating the derivative for any arbitrary set of inputs to $F(\cdot)$. This yields

$$0 = \sum_{n=0}^{\infty} e^{-\rho t}\Bigg\{ \left[u(\boldsymbol{c}_n) - v(N_n) - \frac{\delta}{2}(\pi_n^w)^2 \right]' \frac{d\boldsymbol{g}_t}{d\theta_k} + (\boldsymbol{g}_n \cdot u'(\boldsymbol{c}_n))'\frac{d\boldsymbol{c}_n}{d\theta_k} - (v'(N_n)\mathbf{1} + \delta\pi_n^w\mathbf{1})'\boldsymbol{g}_n\frac{dN_n}{d\theta_k} \Bigg\} dt$$

$$+ \frac{1}{dt}e^{\rho dt}\boldsymbol{\phi}'\frac{d\boldsymbol{V}_0}{d\theta_k} - \frac{1}{dt}e^{\rho dt}\vartheta\frac{d\pi_0^w}{d\theta_k}\frac{1}{d\boldsymbol{x}}$$

where we note that we always have $\frac{dc_{1,n}}{d\theta_k} = 0$ because the planner is constrained by the same boundary condition that the household faces when considering policy perturbations.

Our proof strategy will be to five sets of auxilliary terms to this equation, each of which evaluates to 0, and then use these additional terms to rearrange. In particular, the expressions we add correspond to the discretized competitive equilibrium conditions. And our goal will be to then evaluate the corresponding expression at the stationary equilibrium, group terms, and show that everything evaluates to 0.

Equation 1. We have

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n}\boldsymbol{\phi}'\left[-\rho\boldsymbol{V}_n + \frac{\boldsymbol{V}_{n+1}-\boldsymbol{V}_n}{dt} + u(\boldsymbol{c}_n) - v(N_n) - \frac{\delta}{2}(\pi_n^w)^2 + \boldsymbol{A}^z\boldsymbol{V}_n + \frac{1}{da}(\boldsymbol{s}_n \cdot \boldsymbol{D}_a\boldsymbol{V}_n) \right]$$

where we use $\boldsymbol{\phi} = \boldsymbol{\phi}_{\text{ss}}$. We now use auxilliary results and derivations from before to rewrite this equation as

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n}\boldsymbol{\phi}'\left[u(\boldsymbol{c}_n) - v(N_n) - \frac{\delta}{2}(\pi_n^w)^2 + \boldsymbol{A}^z\boldsymbol{V}_n + \frac{1}{da}(\boldsymbol{s}_n \cdot \boldsymbol{D}_a\boldsymbol{V}_n) \right] - e^{\rho dt}\boldsymbol{\phi}'\frac{1}{dt}\boldsymbol{V}_0$$

Differentiating, we obtain

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \left[(\boldsymbol{\phi} \cdot u'(\boldsymbol{c}_n)) \frac{d\boldsymbol{c}_n}{d\theta_k} - \boldsymbol{\phi}'\mathbf{1} \Big(v'(N_t) \frac{dN_n}{d\theta_k} + \delta \pi_n^w \frac{d\pi_n^w}{d\theta_k} \Big) \right.$$
$$\left. + \boldsymbol{\phi}' \boldsymbol{A}^z \frac{d\boldsymbol{V}_n}{d\theta_k} + \frac{1}{da} (\boldsymbol{\phi} \cdot \boldsymbol{D}_a \boldsymbol{V}_n)' \frac{d\boldsymbol{s}_n}{d\theta_k} + \frac{1}{da} \boldsymbol{\phi}' \boldsymbol{s}_n \cdot \boldsymbol{D}_a \frac{d\boldsymbol{V}_n}{d\theta_k} \right] - e^{\rho dt} \boldsymbol{\phi}' \frac{1}{dt} \frac{d\boldsymbol{V}_0}{d\theta_k}.$$

This is the first auxilliary equation that we will add to our desired expression.

Equation 2. We obtain the second auxilliary condition by simply differentiating the consumption first-order condition. We rewrite the equation as

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \left[\boldsymbol{\chi}'_{[2:J]} u''(\boldsymbol{c}_{n,[2:J]}) - e^{\rho dt} \frac{1}{da} \boldsymbol{V}_n' \boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix} \right] + e^{\rho dt} \frac{1}{da} \boldsymbol{V}_0' \boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix}$$

where we use $\boldsymbol{\chi} = \boldsymbol{\chi}_{\text{ss}}$, and then differentiate to obtain

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \left[\Big(\boldsymbol{\chi}_{[2:J]} \cdot u''(\boldsymbol{c}_{n,[2:J]}) \Big)' \frac{d\boldsymbol{c}_{n,[2:J]}}{d\theta_k} - e^{\rho dt} \frac{1}{da} \left[\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix} \right]' \frac{d\boldsymbol{V}_n}{d\theta_k} \right] + e^{\rho dt} \frac{1}{da} \left[\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix} \right]' \frac{d\boldsymbol{V}_0}{d\theta_k}$$

Equation 3. For our third auxilliary equation, we differentiate the discretized Kolmogorov forward equation. From before, we have

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \left[-\rho \boldsymbol{\lambda}' \boldsymbol{g}_n + \boldsymbol{\lambda}' (\boldsymbol{A}^z)' \boldsymbol{g}_n - \frac{1}{da} (\boldsymbol{s}_n \cdot \boldsymbol{g}_n)' \boldsymbol{D}_a' \boldsymbol{\lambda} \right] + \frac{1}{dt} e^{\rho dt} \boldsymbol{\lambda}' \boldsymbol{g}_0$$

Differentiating with respect to θ_k, we obtain

$$0 = = \sum_{n=0}^{\infty} e^{-\rho t_n} \left[-\rho \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_n}{d\theta_k} + \boldsymbol{\lambda}' (\boldsymbol{A}^z)' \frac{d\boldsymbol{g}_n}{d\theta_k} + (\boldsymbol{g}_n \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{s}_n}{d\theta_k} + (\boldsymbol{s}_n \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{g}_n}{d\theta_k} \right] + \frac{1}{dt} e^{\rho dt} \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_0}{d\theta_k}$$

where we again use $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\text{ss}}$.

Equation 4. We have the aggregate resource constraint with

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \mu \left[\frac{1}{d\boldsymbol{x}} A_n N_n - \boldsymbol{c}_n' \boldsymbol{g}_n \right],$$

where we use $\mu = \mu_{\text{ss}}$. Differentiating, we have

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \mu \left[\frac{1}{d\boldsymbol{x}} A_n \frac{dN_n}{d\theta_k} - \boldsymbol{c}_n' \frac{d\boldsymbol{g}_n}{d\theta_k} - \boldsymbol{g}_n' \frac{d\boldsymbol{c}_n}{d\theta_k} \right].$$

Equation 5. And finally, we use the Phillips curve, which we rewrite using previous results as

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \vartheta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A_n (\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) N_n + \frac{1}{dt} e^{\rho dt} \vartheta \pi_0^w$$

Differentiating, we obtain

$$0 = \sum_{n=0}^{\infty} e^{-\rho t_n} \vartheta \frac{\epsilon}{\delta} \Bigg[\left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A_n \left((\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \frac{d\boldsymbol{g}_n}{d\theta_k} + (\boldsymbol{z} \cdot u''(\boldsymbol{c}_n) \cdot \boldsymbol{g}_n)' \frac{d\boldsymbol{c}_n}{d\theta_k} \right) d\boldsymbol{x} - v''(N_n) \frac{dN_n}{d\theta_n} \right) N_n$$
$$+ \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A_n (\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) \frac{dN_n}{d\theta_k} + \frac{1}{dt} e^{\rho dt} \vartheta \frac{d\pi_0^w}{d\theta_k}$$

Evaluate at stationary Ramsey plan. Crucially, each of our five auxilliary equations must necessarily also hold when evaluated at a stationary Ramsey plan. The key step now, is to evaluate each of the first-order derivatives we taken at the stationary Ramsey plan.

Putting everything together. Having evaluated all derivatives around the stationary Ramsey plan, we add the five auxilliary equations we have derived to the expression for $\frac{dL^{\text{TD}}}{d\theta_k}$ which we started out with, where we now also evaluate the latter at the stationary Ramsey plan. This yields

$$0 = \sum_{n=0}^{\infty} e^{-\rho t} \Bigg\{ \left[u(\boldsymbol{c}) - v(N) - \frac{\delta}{2}(\pi^w)^2 \right]' \frac{d\boldsymbol{g}_t}{d\theta_k} + (\boldsymbol{g} \cdot u'(\boldsymbol{c}))' \frac{d\boldsymbol{c}_n}{d\theta_k} - (v'(N)\mathbf{1} + \delta\pi^w \mathbf{1})' \boldsymbol{g} \frac{dN_n}{d\theta_k}$$

$$+ (\boldsymbol{\phi} \cdot u'(\boldsymbol{c})) \frac{d\boldsymbol{c}_n}{d\theta_k} - \boldsymbol{\phi}' \mathbf{1} \left(v'(N) \frac{dN_n}{d\theta_k} + \delta \pi_n^w \frac{d\pi_n^w}{d\theta_k} \right)$$

$$+ \boldsymbol{\phi}' \boldsymbol{A}^z \frac{d\boldsymbol{V}_n}{d\theta_k} + \frac{1}{da} (\boldsymbol{\phi} \cdot \boldsymbol{D}_a \boldsymbol{V})' \frac{d\boldsymbol{s}_n}{d\theta_k} + \frac{1}{da} \boldsymbol{\phi}' \boldsymbol{s} \cdot \boldsymbol{D}_a \frac{d\boldsymbol{V}_n}{d\theta_k}$$

$$+ \left(\boldsymbol{\chi}_{[2:J]} \cdot u''(\boldsymbol{c}_{[2:J]}) \right)' \frac{d\boldsymbol{c}_{n,[2:J]}}{d\theta_k} - e^{\rho dt} \frac{1}{da} \left[\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix} \right]' \frac{d\boldsymbol{V}_n}{d\theta_k}$$

$$- \rho \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_n}{d\theta_k} + \boldsymbol{\lambda}' (\boldsymbol{A}^z)' \frac{d\boldsymbol{g}_n}{d\theta_k} + (\boldsymbol{g} \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{s}_n}{d\theta_k} + (\boldsymbol{s} \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{g}_n}{d\theta_k}$$

$$+ \mu \frac{1}{d\boldsymbol{x}} A \frac{dN_n}{d\theta_k} - \mu \boldsymbol{c}' \frac{d\boldsymbol{g}_n}{d\theta_k} - \mu \boldsymbol{g}' \frac{d\boldsymbol{c}_n}{d\theta_k}$$

$$+ \vartheta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A \left((\boldsymbol{z} \cdot u'(\boldsymbol{c}))' \frac{d\boldsymbol{g}_n}{d\theta_k} + (\boldsymbol{z} \cdot u''(\boldsymbol{c}) \cdot \boldsymbol{g})' \frac{d\boldsymbol{c}_n}{d\theta_k} \right) d\boldsymbol{x} - v''(N) \frac{dN_n}{d\theta_n} \right) N_n$$

$$+ \vartheta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A (\boldsymbol{z} \cdot u'(\boldsymbol{c}))' \boldsymbol{g} d\boldsymbol{x} - v'(N) \right) \frac{dN_n}{d\theta_k} \Bigg\} dt$$

$$+ e^{\rho dt} \boldsymbol{\phi}' \frac{d\boldsymbol{V}_0}{d\theta_k} - e^{\rho dt} \vartheta \frac{d\pi_0^w}{d\theta_k} \frac{1}{d\boldsymbol{x}}$$

$$- e^{\rho dt} \boldsymbol{\phi}' \frac{d\boldsymbol{V}_0}{d\theta_k} + e^{\rho dt} \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_0}{d\theta_k} + e^{\rho dt} \vartheta \frac{d\pi_0^w}{d\theta_k} \frac{1}{d\boldsymbol{x}}$$

$$+ dt e^{\rho dt} \frac{1}{da} \left[\boldsymbol{D}_a' \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{[2:J]} \end{pmatrix} \right]' \frac{d\boldsymbol{V}_0}{d\theta_k}$$

where every term that does not have a time step subscript n is understood to have been evaluated at the stationary Ramsey plan.

Our proof is now almost complete. First, note how the timeless penalties *exactly offset* the "boundary terms" that resulted from rearranging the forward looking implementability conditions. In particular, notice that $\frac{d\boldsymbol{g}_0}{d\theta_k} = 0$ and the term in the very last line goes to 0 as $dt \to 0$. The remaining boundary (or initial condition) terms exactly cancel out.

Second, we plug in for

$$\frac{d\boldsymbol{s}_n}{d\theta_k} = \frac{dr_n}{d\theta_k} a + zw \frac{dN_n}{d\theta_k} + zN \frac{dw_n}{d\theta_k} - \frac{d\boldsymbol{c}_n}{d\theta_k}$$

when evaluated at the stationary Ramsey plan.

Third and finally, we group all terms by *derivatives*. After this last step, we see that the grouped expressions correspond *exactly* to the optimality conditions that define the stationary Ramsey plan. Consequently, they must be 0. This concludes the proof: We started with an expression for $\frac{dL^{\text{TD}}}{d\theta_k}$, and added five auxilliary expressions, each of which itself evaluated to 0. Then we evaluated the resulting expression around the stationary Ramsey plan and showed that it was 0. Consequently, we have shown that

$$\frac{dL^{\text{TD}}}{d\theta_k} = 0$$

when evaluated at the stationary Ramsey plan. And since k was arbitrary, we have our desired result for any policy perturbation around the stationary Ramsey plan. We have thus shown that Ramsey policy according to the timeless dual Lagrangian L^{TD} is indeed time consistent.

A.8 The Timeless Ramsey Problem in Dual Form

The timeless Ramsey problem also admits a dual representation, which we introduce next. The distinction between the primal and dual problems lies in the treatment of the constraints that a planner faces. In the primal approach, the planner optimizes over allocations, prices, and instruments given a set of constraints or implementability conditions. In the dual approach, the planner explicitly optimizes over the policy instrument, in this case, interest rates, using the implementability conditions to characterize the comparative statics of endogenous variables to policy.[42] The primal and dual representations of the timeless Ramsey problem have their distinct advantages and we leverage both in our analysis.

Definition. (**Timeless Dual Ramsey Problem**) *A timeless dual Ramsey problem solves*

$$\max_{\{i_t\}} L^{\mathrm{TD}}(g_0(\cdot), \phi(\cdot), \vartheta),$$

where $L^{\mathrm{TD}}(g_0(\cdot), \phi(\cdot), \vartheta)$ denotes the timeless dual Lagrangian, given an initial distribution $g_0(\cdot)$ as well as initial promises $\phi(\cdot)$ and ϑ. The Lagrangian is defined as

$$L^{\mathrm{TD}}(g_0(\cdot), \phi(\cdot), \vartheta) = \int_0^\infty e^{-\rho t} \iint \left[u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 \right] g_t(a,z)\, da\, dz\, dt + \mathcal{T}(\phi(\cdot), \vartheta), \quad (56)$$

where all endogenous variables are understood as functions of the policy path $\{i_t\}$.

Proposition 23. (**Time-Consistency of the Timeless Ramsey Problem**) *Optimal policy under the timeless primal and dual Lagrangians is time-consistent around the stationary Ramsey plan under the timeless penalty $\mathcal{T}(\phi_{ss}(\cdot), \vartheta_{ss})$. That is,*

$$\frac{d}{d\boldsymbol{\theta}} L^{\mathrm{TD}}\Big(g_{\mathrm{ss}}(\cdot), \phi_{\mathrm{ss}}(\cdot), \vartheta_{\mathrm{ss}}, \boldsymbol{\theta}_{\mathrm{ss}}, \boldsymbol{Z}_{\mathrm{ss}} \Big) = 0. \quad (57)$$

[42] In simple terms, a useful analogy may be to interpret the dual approach as substituting constraints into the objective of an optimization problem, and the primal approach as accounting for constraints as additional terms in a Lagrangian.

A.9 Sequence Space Perturbations in the Dual

In this section, we develop a sequence-space perturbation approach to solve optimal stabilization policy in the dual. We take as our starting point not equation (51) but a sequence-space representation of the timeless dual Lagrangian, which we now introduce.

The timeless dual Lagrangian defined in Appendix A.8 takes as its inputs (i) the time paths of allocations and prices, (ii) an initial distribution, and (iii) initial timeless penalties. Unlike in the primal form, it does not explicitly feature the time paths of multipliers. For a given path of policy $\boldsymbol{i}$ and shocks $\boldsymbol{Z}$, we can directly use the equilibrium map (50) to solve out for allocations and prices, i.e., $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z})$. A sequence-space representation of the timeless dual Lagrangian is then given by

$$L^{\mathrm{TD}}(\boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z}), \boldsymbol{i}, \boldsymbol{Z}), \tag{58}$$

where we again leave implicit the dependence of $L^{\mathrm{TD}}(\cdot)$ on $g_0(a, z)$ as well as $\phi(a, z)$ and θ.

The timeless dual Lagrangian (58) implies the local efficiency criterion for optimal policy

$$\mathcal{F}(\boldsymbol{i}, \boldsymbol{Z}) = \frac{d}{d\boldsymbol{i}} L^{\mathrm{TD}}(\boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z}), \boldsymbol{i}, \boldsymbol{Z}) = 0, \tag{59}$$

where $\mathcal{F}(\cdot)$ implicitly takes as given an initial distribution as well as initial promises. Equation (59) represents the planner's necessary first-order optimality condition in sequence-space form. We can use it to directly characterize optimal policy in the dual in terms of exogenous shocks, i.e.,

$$\boldsymbol{i} = \boldsymbol{i}(\boldsymbol{Z}).$$

Importantly, given policy $\boldsymbol{i}$, the timeless penalty (ϕ, ϑ) does not affect competitive equilibrium, summarized by (50). It only influences the planner's assessment of optimal policy.

Proposition 24. (**Optimal Policy Perturbations in the Dual**) *Consider the dual Ramsey problem, under which a locally efficient policy is characterized by $\mathcal{F}(\cdot) = 0$. Suppose we initialize the Ramsey problem at the stationary Ramsey plan, with $g_0(a, z) = g_{\mathrm{ss}}(a, z)$, and with initial timeless penalties $\phi(a, z) = \phi_{\mathrm{ss}}(a, z)$ and $\theta = \theta_{\mathrm{ss}}$. To first order, optimal stabilization policy is then characterized by*

$$d\boldsymbol{i} = -\mathcal{F}_{\boldsymbol{i}}^{-1} \mathcal{F}_{\boldsymbol{Z}} d\boldsymbol{Z}, \tag{60}$$

where $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{\mathrm{ss}}$ is the exogenous shock, and where $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{Z}}$ denote Hessians of the timeless dual Lagrangian.

We prove Proposition 24 in Appendix B.4.

In the dual, approximating optimal policy using sequence-space perturbation methods requires computing second-order total derivatives of the timeless dual Lagrangian. These are in turn given by the Jacobians of the planner's first-order condition, i.e., $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{Z}}$. The recent literature on perturbation methods in heterogeneous-agent economies has shown that sequence-space Jacobians, i.e., first-order derivatives of model objects in sequence-space representation, are sufficient to characterize transition dynamics to first order. Similarly, we have shown in Section 5.1.2 that computing optimal policy and Ramsey plans in the primal also only requires sequence-space Jacobians. Computing optimal policy and welfare in the dual representation of our Ramsey problem requires a second-order analysis, however.

To that end, we introduce *sequence-space Hessians* as the natural, second-order generalization of sequence-space Jacobians. In Appendix B.5, we formally define sequence-space Hessians, and we show both how to efficiently compute and leverage them to characterize optimal policy in the dual. We extend the methodology developed by Auclert et al. (2021) to problems that require second-order derivatives, i.e., sequence-space Hessians. While we focus on their use in the context of computing Ramsey plans in the dual, we argue that sequence-space Hessians are useful more broadly whenever a second-order analysis is required.

A.9.1 Advantages and Disadvantages of Primal and Dual Formulations

Our approach allows for a representation of Ramsey problems in either the primal or the dual form, and we show how to characterize and compute optimal policy in both cases. We conclude this section with a discussion of the advantages and disadvantages of both approaches.

The primal form and, in particular, the associated Ramsey plan representation (51) are more conducive to computing optimal policy non-linearly. Computing the Ramsey plan in the primal requires solving a system of non-linear equations. Fully optimized quasi-Newton methods and other non-linear equation solvers can be leveraged for this task.

In the primal approach, however, we have to compute multipliers and their transition dynamics. In the Ramsey plan representation (51), multipliers $\boldsymbol{M}$ enter as an explicit argument. It is not generically possible to characterize Ramsey plans in the spirit of (51) without multipliers. The dual approach, on the other hand, takes as its starting point the timeless dual Lagrangian, which does not explicitly depend on multipliers. Consequently, it is not necessary to compute the time paths of multipliers to characterize optimal policy in the dual.

The relative disadvantage of the dual approach, however, is that a first-order approximation of optimal policy requires second-order total derivatives of the timeless dual Lagrangian. Unlike in the primal approach, which only requires computing sequence-space Jacobians, in the dual we have to compute sequence-space Hessians. Computationally, this is a more complex task both in terms of compute time and memory demands. In summary, therefore, the main tradeoff between the primal and dual approaches is that the former requires computing the time paths of multipliers, while the latter requires sequence-space Hessians instead of only Jacobians.

Another advantage of the dual representation is that it provides an easily implementable local efficiency criterion. In particular, assessing whether a policy $\boldsymbol{i}$ is locally efficient in the dual only requires computing first-order derivatives of the timeless dual Lagrangian, which is possible in terms of sequence-space Jacobians. Unlike in the primal, however, efficiency assessments in the dual do not require computing the derivatives of multipliers. In practice, even if optimal policy is computed in the primal, verifying local efficiency in the dual is a cheap but helpful exercise.

Lastly, sequence-space Hessians and, more broadly, second-order sequence-space perturbation methods likely have many useful applications beyond computing Ramsey plans in the dual. We therefore view our treatment of sequence-space Hessians as a standalone contribution of this paper.

B Optimal Policy and Ramsey Plans in Sequence Space

In this Appendix, we discuss how to operationalize our method and compute optimal policy numerically. Following much of the recent literature on computational methods in heterogeneous-agent economies, we work with a sequence-space representation of our model. In the interest of accessibility, we follow the notation and conventions of Auclert et al. (2021) as closely as possible, extending their work on sequence-space Jacobians to Ramsey problems and welfare analysis. While they work in discrete time, we show below that continuous-time heterogeneous-agent models are nested by the same general model representation they propose. To establish this relationship, we first discretize our model following the same steps that would also be required for numerical implementation.

Discretization. We first discretize the equations that characterize competitive equilibrium and optimal policy in both time and space. We use a finite-difference discretization scheme building on Achdou et al. (2021).[43] In particular, we discretize the time dimension over a finite horizon, $t \in [0, T]$ where T can be arbitrarily large, using N discrete time steps, which we denote by $n = 1, \ldots, N$. With a step size $dt = \frac{T}{N-1}$, we have $t_n = dt(n-1)$. We similarly discretize the idiosyncratic state space over (a, z) using J grid points. Using bold-faced notation, we denote the discretized consumption policy function of the household at time t_n as the $J \times 1$ vector $\boldsymbol{c}_n$, where the ith element corresponds to $c_{t_n}(a_i, z_i)$.

B.1 Sequence-Space Representation of Equilibrium

After discretizing our model, the resulting equations satisfy the general model representation of heterogeneous-agent economies presented in Auclert et al. (2021). To facilitate comparison, we follow their notation in this Appendix. We consider a general representation of a heterogeneous-agent problem as a mapping from time paths of aggregate inputs $(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Z})$ to time paths of aggregate outputs $\boldsymbol{Y}$. We use bold-faced notation here to indicate time paths, with $\boldsymbol{\theta} = \{\theta_n\}_{n=1}^{N}$. It will be useful to explicitly distinguish between the time paths for policy $\boldsymbol{\theta}$ and the exogenous shock $\boldsymbol{Z}$ on the one hand, and the time paths for other aggregate inputs $\boldsymbol{X}$ on the other hand. To simplify the exposition, we assume that there is only one aggregate input variable other than policy and the shock, so that $X_n \in \mathbb{R}$.

Denoting the discretized cross-sectional distribution by the $J \times 1$ vector $\boldsymbol{g}_n$, our main focus will be on outcome variables that take the form $Y_n = \boldsymbol{y}_n' \boldsymbol{g}_n$, where $\boldsymbol{y}_n$ is a $J \times 1$ vector that represents an individual outcome.[44] For example, aggregate consumption takes the form $C_n = \boldsymbol{c}_n' \boldsymbol{g}$. Given an

[43] For a detailed description of the discretization procedure, see Achdou et al. (2021) or Schaab and Zhang (2021). We also leverage the adaptive sparse grid method developed by Schaab and Zhang (2021) and Schaab (2020) to solve dynamic programming problems in continuous time.

[44] We normalize the discretized distribution representation so that $\boldsymbol{g}_n$ sums to 1, i.e., $\mathbf{1}' \boldsymbol{g}_n = 1$, where $\mathbf{1}$ is a $J \times 1$ vector of 1s.

initial distribution $\boldsymbol{g}_0$, aggregate outcomes $\boldsymbol{Y}$ then solve the system of equations

$$\boldsymbol{V}_n = \nu(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n) \tag{61}$$

$$\boldsymbol{g}_{n+1} = \Lambda(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n)\boldsymbol{g}_n \tag{62}$$

$$Y_n = y(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n)'\boldsymbol{g}_n. \tag{63}$$

The implementability conditions of our baseline HANK economy can be expressed in terms of the time paths of macroeconomic aggregates as well as those of aggregate outcomes $\boldsymbol{Y}$. Using the above representation, aggregate outcomes $\boldsymbol{Y}$ can in turn be expressed in terms of the time paths of aggregate allocations and prices $\boldsymbol{X}$, policy $\boldsymbol{\theta}$, and shocks $\boldsymbol{Z}$. In summary, equilibria given policy and shocks can be expressed in terms of the *equilibrium map* $\mathcal{H}(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Z}) = 0$.

B.2 Sequence-Space Representation of Ramsey Plans

We now show how to express the Ramsey plan optimality conditions, which characterize the multipliers and optimal policy, in a general model representation akin to equations (61) through (63). In general, Ramsey plans in heterogeneous-agent economies feature three types of multipliers: aggregate multipliers, individual forward-looking multipliers, and individual backward-looking multipliers. In our baseline environment, the aggregate multipliers are ϑ_t and μ_t, the individual forward-looking multiplier is $\lambda_t(a,z)$, and the (system of) individual backward-looking multipliers is $\phi_t(a,z)$ and $\chi_t(a,z)$.

The Ramsey plan representation (51) summarizes the optimality conditions of the timeless Ramsey plan in sequence-space form. In particular, the Ramsey map $\mathcal{R}(\cdot)$ takes the time paths of aggregate multipliers $\boldsymbol{M}$ as explicit inputs. Our goal now is to show that the optimality conditions of the timeless Ramsey plan can be written in terms of $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{\theta})$ and $\boldsymbol{Z}$.

Forward-looking individual multipliers take the form

$$\boldsymbol{\lambda}_n = f(\boldsymbol{\lambda}_{n+1}, \boldsymbol{V}_n, X_n, M_n, \theta_n, Z_n), \tag{64}$$

which is analogous to equation (61), which characterizes individual forward-looking behavior. For example, it is straightforward to verify that equation (34) satisfies this form: it expresses today's multiplier $\lambda_n(a,z)$ in terms of today's aggregate multipliers, individual allocations, aggregate allocations and prices, as well as the future multiplier $\lambda_{n+1}(a,z)$.

Analogously to equation (61), the recursive structure of forward-looking individual multipliers allows us to efficiently compute their first-order derivatives. We summarize this observation in the following Lemma.

Lemma 25. *For any $k \geq 1$, we have*

$$\frac{\partial \boldsymbol{\lambda}_n}{\partial \theta_k} = \begin{cases} 0 & \text{if} \quad n > k \\ \frac{\partial \boldsymbol{\lambda}_{n-s}}{\partial \theta_{k-s}} & \text{else for } s < n \end{cases}$$

and likewise for first-order derivatives in X_k, M_n, and Z_n.

Backward-looking individual multipliers typically correspond to promises that the Ramsey planner makes to individuals. They are characterized by a particular kind of Kolmogorov forward equation. In particular, promise-keeping Kolmogorov forward equations feature a forcing term that captures the "births" and "deaths" of promises, captured by $\partial_a \chi_t(a, z)$ in equation (33). Consequently, backward-looking multipliers can be represented as

$$\boldsymbol{\phi}_{n+1} = \Lambda(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n)\boldsymbol{\phi}_n + b(\boldsymbol{\phi}_n, \boldsymbol{\lambda}_n, \boldsymbol{V}_n, \boldsymbol{g}_n, X_n, M_n, \theta_n, Z_n). \tag{65}$$

The multiplier representations (64) and (65), together with equations (61) through (63) and the equilibrium map (50), let us conclude that timeless Ramsey plans admit the sequence-space representation (51).

B.3 Proof of Proposition 12

Denote the Ramsey plan by $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{\theta})$. Using a first-order Taylor expansion around the stationary Ramsey plan, we have

$$\mathcal{R}(\boldsymbol{R}, \boldsymbol{Z}) \approx \mathcal{R}(\boldsymbol{R}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}) + \mathcal{R}_{\boldsymbol{R}}(\boldsymbol{R}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}})d\boldsymbol{R} + \mathcal{R}_{\boldsymbol{Z}}(\boldsymbol{R}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}})d\boldsymbol{Z}.$$

Notice that we have $\mathcal{R}(\boldsymbol{R}, \boldsymbol{Z}) = 0$ by the definition of $\boldsymbol{R}$ as a Ramsey plan, i.e., as solving $\mathcal{R}(\cdot) = 0$ for a given $\boldsymbol{Z}$ as in (51).

We now show that $\mathcal{R}(\boldsymbol{R}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}) = 0$ as well, assuming, as we do in Proposition 12, that the Ramsey problem is initialized at $(g_0(\cdot), \phi(\cdot), \vartheta) = (g_{\text{ss}}(\cdot), \phi_{\text{ss}}(\cdot), \vartheta_{\text{ss}})$. The Ramsey plan map $\mathcal{R}(\cdot)$ is a system of equations that comprises two sets of conditions, those for competitive equilibrium as well as the first-order conditions associated with the timeless Ramsey problem in the primal representation. By definition, the stationary Ramsey plan comprises a feasible competitive equilibrium. Consequently, when evaluated at $(\boldsymbol{R}, \boldsymbol{Z}) = (\boldsymbol{R}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}})$, those conditions in $\mathcal{R}(\cdot)$ associated with competitive equilibrium are 0.

That leaves the first-order conditions associated with the timeless primal Ramsey problem. It follows from Proposition 6 that the timeless primal Ramsey problem is time-consistent, so that the planner does not want to deviate from the stationary Ramsey plan when $\boldsymbol{Z} = \boldsymbol{Z}_{\text{ss}}$. It also follows from our duality proof for the primal and dual representations that $\frac{d}{d\boldsymbol{\theta}} L^{\text{TP}}(\boldsymbol{\theta}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}) = 0$ also implies that each of the associated first-order conditions of the timeless primal problem are 0 when evaluated at $(\boldsymbol{R}_{ss}, \boldsymbol{Z}_{ss})$. Putting these observations together implies $\mathcal{R}(\boldsymbol{R}_{ss}, \boldsymbol{Z}_{ss}) = 0$.

We are then simply left with

$$0 \approx \mathcal{R}_{\boldsymbol{R}}(\boldsymbol{R}_{ss}, \boldsymbol{Z}_{ss}) d\boldsymbol{R} + \mathcal{R}_{\boldsymbol{Z}}(\boldsymbol{R}_{ss}, \boldsymbol{Z}_{ss}) d\boldsymbol{Z}.$$

Rearranging and inverting $\mathcal{R}_{\boldsymbol{R}}$ yields the desired result.

B.4 Proof of Proposition 24

We proceed as follows. A first-order Taylor approximation of $F(\cdot)$ (in $\boldsymbol{\theta}$ and $\boldsymbol{Z}$) around the stationary Ramsey plan yields

$$\begin{aligned} F(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}, \boldsymbol{Z}) =& F(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}_{ss}, \boldsymbol{Z}_{ss}) \\ &+ F_{\boldsymbol{\theta}}(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}_{ss}, \boldsymbol{Z}_{ss})(\boldsymbol{\theta} - \boldsymbol{\theta}_{ss}) + F_{\boldsymbol{Z}}(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}_{ss}, \boldsymbol{Z}_{ss})(\boldsymbol{Z} - \boldsymbol{Z}_{ss}). \end{aligned}$$

First, we must have

$$F(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}, \boldsymbol{Z}) = 0$$

by construction because that's our definition for optimal policy $\boldsymbol{\theta}(\boldsymbol{Z})$. Second, we also have

$$F(\boldsymbol{g}_{ss}, \boldsymbol{\phi}_{ss}, \vartheta_{ss}, \boldsymbol{\theta}_{ss}, \boldsymbol{Z}_{ss}) = 0,$$

which is the main result of Section 4.5 and whose proof was just presented in the previous Appendix subsection. Denoting $d\boldsymbol{\theta} = \boldsymbol{\theta} - \boldsymbol{\theta}_{ss}$ and $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{ss}$, we thus have

$$0 = F_{\boldsymbol{\theta}} d\boldsymbol{\theta} + F_{\boldsymbol{Z}} d\boldsymbol{Z},$$

where the Jacobians of $F(\cdot)$ are evaluated at the stationary Ramsey plan.

B.5 Sequence-Space Hessians

To compute optimal policy to first order using Proposition 24, we effectively need to differentiate $L^{\text{TD}}(\cdot)$ twice. In particular, $F(\cdot) = \frac{d}{d\boldsymbol{\theta}} L^{\text{TD}}$ features first-order derivative terms, which can be cast as sequence-space Jacobians (Auclert et al., 2021). Therefore, computing the total derivatives $\frac{d}{d\boldsymbol{\theta}} F(\cdot)$ and $\frac{d}{d\boldsymbol{Z}} F(\cdot)$, which are used in equation (60) to characterize optimal policy $d\boldsymbol{\theta}$, we require second-order derivatives. Consequently, computing optimal stabilization policy using our approach requires that we compute both first- and second-order total derivatives of all objects that feature in the timeless dual Lagrangian.[45]

[45] An alternative approach is to work in the primal representation of our timeless Ramsey problem and solve for the time-consistent Ramsey plan—including all multipliers—to first order using the system of equations that characterizes the Ramsey plan. We develop this alternative approach in ongoing work. The dual approach avoids having to compute the multipliers but requires sequence-space Hessians. For the primal approach, we have to compute an extended set

In a sequence-space representation of our model, these objects are all functions of the time paths of aggregate inputs, i.e., $(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Z})$, where $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{\theta}, \boldsymbol{Z})$. Moreover, the timeless dual Lagrangian itself can be represented in terms of aggregate outcomes $\boldsymbol{Y}$, using the general model representation above. Consequently, computing the matrices $\frac{d}{d\boldsymbol{\theta}}F$ and $\frac{d}{d\boldsymbol{Z}}F$ requires taking total derivatives of specific aggregate outcomes $\boldsymbol{Y}$.

We define *sequence-space Hessians* as the matrices of mixed partial derivatives of model objects that can be represented as functions of aggregate sequences around the stationary Ramsey plan. We discuss these mixed partial derivative matrices in detail in Section B.5.1. Subsequently, in Section B.5.2, we show how to build up the second-order total derivatives of the timeless dual Lagrangian, i.e., $\frac{d}{d\boldsymbol{\theta}}F$ and $\frac{d}{d\boldsymbol{Z}}F$, from sequence-space Hessians.

B.5.1 A Fake-News Algorithm to Compute Sequence-Space Hessians

We now extend the fake-news algorithm of Auclert et al. (2021) to compute sequence-space Hessians, i.e., the matrices of mixed partial derivatives $\frac{\partial^2}{\partial\theta_k\partial\theta_l}Y_n$ in a sequence-space representation of the model. The results we present below hold for any mixed partial derivative of $Y_n(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Z})$, but to ease notation we focus specifically on the mixed derivative $\frac{\partial^2}{\partial\theta_k\partial\theta_l}$ for some given $k, l \in \{1, \ldots, N\}$.

Using equation (63), we can rewrite the mixed derivative of aggregate outcome Y_n at time t_n as

$$\frac{\partial^2 Y_n}{\partial\theta_k\partial\theta_l} = \frac{\partial}{\partial\theta_l}\left(\boldsymbol{y}_n' \frac{\partial \boldsymbol{g}_n}{\partial\theta_k} + \boldsymbol{g}_n' \frac{\partial \boldsymbol{y}_n}{\partial\theta_k}\right) = \boldsymbol{y}_n' \frac{\partial^2 \boldsymbol{g}_n}{\partial\theta_k\partial\theta_l} + \frac{\partial \boldsymbol{g}_n}{\partial\theta_k}' \frac{\partial \boldsymbol{y}_n}{\partial\theta_l} + \boldsymbol{g}_n' \frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k\partial\theta_l} + \frac{\partial \boldsymbol{y}_n}{\partial\theta_k}' \frac{\partial \boldsymbol{g}_n}{\partial\theta_l}$$

This derivation underscores that we generally need both the first-order and second-order mixed partial derivatives of individual outcomes $\boldsymbol{y}_n$ and the distribution $\boldsymbol{g}_n$ to compute aggregate sequence-space Hessians $\partial^2\boldsymbol{Y}$. Our method leverages several useful properties of these first- and second-order derivatives. In the following, we prove key properties of the second-order mixed derivatives $\frac{\partial^2}{\partial\theta_k\partial\theta_l}\boldsymbol{y}_n$ and $\frac{\partial^2}{\partial\theta_k\partial\theta_l}\boldsymbol{g}_n$, and we refer the reader to Auclert et al. (2021) for the properties of the first-order partial derivatives.

First, notice that mixed partial derivatives are symmetric, or interchangeable, by the standard continuity argument. That is

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k\partial\theta_l} = \frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_l\partial\theta_k} \quad \text{and} \quad \frac{\partial^2 \boldsymbol{g}_n}{\partial\theta_k\partial\theta_l} = \frac{\partial^2 \boldsymbol{g}_n}{\partial\theta_l\partial\theta_k}.$$

Second, the recursive structure of the system (61) - (63) gives rise to the following key property of mixed partial derivatives in sequence space.

---

of sequence-space Jacobians, solving to first order for the transition paths of i) allocations, ii) optimal policy, and iii) multipliers.

Lemma 26. *We have*

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k \partial\theta_l} = \begin{cases} 0 & \text{if} \quad n > \min\{k,l\} \\ \frac{\partial^2 \boldsymbol{y}_{n-s}}{\partial\theta_{k-s}\partial\theta_{l-s}} & \text{else for } s < n \end{cases}$$

Leveraging these first two properties of mixed derivatives of individual outcomes in sequence space, we can construct sequence-space Hessian matrices using the following shortcut: Instead of computing all N^2 numerical derivatives, we simply compute

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k \partial\theta_N}$$

for $1 \leq k \leq N$, which requires only N numerical derivative evaluations.[46]

Third, we exploit the fact that the transition matrix Λ, which describes the law of motion of the cross-sectional distribution in equation (63), has a particular structure in continuous time. In particular, we have

$$\Lambda(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n) = 1 + dt\boldsymbol{A}(\boldsymbol{V}_{n+1}, X_n, \theta_n, Z_n)',$$

where $\boldsymbol{A}$ is the $J \times J$ matrix that discretizes the HJB operator $\mathcal{A}$, and $\boldsymbol{A}'$, its transpose, is the analog for the adjoint $\mathcal{A}^*$. In our baseline HANK model, the discretized transition matrix takes the form

$$\boldsymbol{A}_n = \boldsymbol{s}_n \cdot \boldsymbol{D}_a + \boldsymbol{A}^z \tag{66}$$

where $\boldsymbol{A}^z$ is given exogenously, and its derivatives with respect to θ_k, X_k and Z_k are therefore 0. The matrix $\boldsymbol{D}_a$ discretizes the partial derivative ∂_a as we show in Appendix A, and it is also invariant to perturbations in aggregate inputs as long as the step size used for the numerical derivative is fine enough. In particular, the key insight here is that taking derivatives of the general transition matrix Λ in equation (63) simply amounts to differentiating $\boldsymbol{s}_n$ in equation (66).[47] We record this observation in the following Lemma.

Lemma 27. *The first- and second-order mixed partial derivatives of the transition matrix Λ_n in our setting are given by*

$$\frac{\partial \Lambda_n}{\partial\theta_k} = dt\frac{\partial \boldsymbol{s}_n}{\partial\theta_k} \cdot \boldsymbol{D}_a \quad \text{and} \quad \frac{\partial^2 \Lambda_n}{\partial\theta_k\partial\theta_l} = dt\frac{\partial^2 \boldsymbol{s}_n}{\partial\theta_k\partial\theta_l} \cdot \boldsymbol{D}_a.$$

Fourth, we characterize the properties of the mixed derivatives of the cross-sectional distribution. We assume for simplicity that the economy is initialized at the cross-sectional distribution that

[46] For other mixed derivatives, such as $\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k \partial Z_l}$, we require $2N$ evaluations, i.e., both $\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_k \partial Z_N}$ and $\frac{\partial^2 \boldsymbol{y}_n}{\partial\theta_N \partial Z_l}$.

[47] Notice that equation (66) is specific to our baseline model and consequently breaks with the spirit of generality otherwise adopted in this section. However, an equation *like* (66) will generally hold in any continuous-time heterogeneous-agent model. We think there is some value to highlighting how to leverage this equation when constructing sequence-space Jacobians and Hessians, and we therefore use equation (66) in the following while otherwise maintaining our general notation.

corresponds to the stationary Ramsey plan, that is, $g_1 = g_{ss}$, where we recall that n starts at 1 and $t_1 = 0$. The initial distribution is given exogenously and does not adjust on impact. That is, $\frac{\partial^2 g_1}{\partial \theta_k \partial \theta_l} = 0$. Using equation (63) and Lemma 27, the response of the cross-sectional distribution at time step $n = 2$ is thus

$$\frac{\partial^2 \boldsymbol{g}_2}{\partial \theta_k \partial \theta_l} = \frac{\partial^2 \Lambda_1}{\partial \theta_k \partial \theta_l} \boldsymbol{g}_{ss} = dt \left(\frac{\partial^2 \boldsymbol{s}_1}{\partial \theta_k \partial \theta_l} \cdot \boldsymbol{D}_a \right) \boldsymbol{g}_{ss}$$

We now exploit the recursive structure of equation (63) to derive two alternative expressions for the mixed derivatives $\frac{\partial^2 g_n}{\partial \theta_k \partial \theta_l}$, for $n \geq 3$. We summarize in the next Lemma.

Lemma 28. *The mixed partial derivatives of the cross-sectional distribution $\boldsymbol{g}_n$ at time steps $n \geq 3$ can be computed recursively using*

$$\begin{aligned}\frac{\partial^2 \boldsymbol{g}_n}{\partial \theta_k \partial \theta_l} =& \Lambda_{ss} \frac{\partial^2 \boldsymbol{g}_{n-1}}{\partial \theta_k \partial \theta_l} + \frac{\partial^2 \boldsymbol{g}_2}{\partial \theta_{k-(n-2)} \partial \theta_{l-(n-2)}} \mathbb{1}_{\min\{k-(n-2), l-(n-2)\} \geq 1} \\ &+ dt \left(\frac{\partial \boldsymbol{s}_1}{\partial \theta_{l-(n-2)}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{n-1}}{\partial \theta_k} \mathbb{1}_{l-(n-2) \geq 1} + dt \left(\frac{\partial \boldsymbol{s}_1}{\partial \theta_{k-(n-2)}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{n-1}}{\partial \theta_l} \mathbb{1}_{k-(n-2) \geq 1}\end{aligned}$$

or non-recursively using

$$\begin{aligned}\frac{\partial^2 \boldsymbol{g}_n}{\partial \theta_k \partial \theta_l} =& \sum_{r=1}^{R_1} (\Lambda_{ss})^{n-r-1} \frac{\partial^2 \boldsymbol{g}_2}{\partial \theta_k \partial \theta_l} \\ &+ dt \sum_{r=1}^{R_2} (\Lambda_{ss})^{n-r-2} \left(\frac{\partial \boldsymbol{s}_1}{\partial \theta_{k-r}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{1+r}}{\partial \theta_l} + dt \sum_{r=1}^{R_3} (\Lambda_{ss})^{n-r-2} \left(\frac{\partial \boldsymbol{s}_1}{\partial \theta_{l-r}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{1+r}}{\partial \theta_k}\end{aligned}$$

where $R_1 = \min\{k, l, n-1\}$, $R_2 = \min\{k-1, n-2\}$, and $R_3 = \min\{l-1, n-2\}$.

Fifth and finally, we discuss how to efficiently compute a given mixed partial derivative numerically. The most popular finite-difference stencil to compute second-order mixed derivatives is given by

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial \theta_k \partial \theta_l} = \frac{\boldsymbol{y}_n^{++} - \boldsymbol{y}_n^{+-} - \boldsymbol{y}_n^{-+} + \boldsymbol{y}_n^{--}}{4h^2} \tag{67}$$

where $\boldsymbol{y}_n^{++} = \boldsymbol{y}_n(\ldots, \theta_k + h, \ldots, \theta_l + h, \ldots)$, $\boldsymbol{y}_n^{+-} = \boldsymbol{y}_n(\ldots, \theta_k + h, \ldots, \theta_l - h, \ldots)$, $\boldsymbol{y}_n^{-+} = \boldsymbol{y}_n(\ldots, \theta_k - h, \ldots, \theta_l + h, \ldots)$, and $\boldsymbol{y}_n^{--} = \boldsymbol{y}_n(\ldots, \theta_k - h, \ldots, \theta_l - h, \ldots)$. This stencil requires 4 function evaluations for every mixed derivative and is therefore very costly.

An alternative and, in our case, substantially more efficient stencil is

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial \theta_k \partial \theta_l} = \frac{\boldsymbol{y}_n^{++} - \boldsymbol{y}_n^{+\cdot} - \boldsymbol{y}_n^{\cdot+} + 2\boldsymbol{y}_n - \boldsymbol{y}_n^{-\cdot} - \boldsymbol{y}_n^{\cdot-} + \boldsymbol{y}_n^{--}}{2h^2} \tag{68}$$

where $\boldsymbol{y}_n^{+\cdot} = \boldsymbol{y}_n(\ldots, \theta_k + h, \ldots, \theta_l, \ldots)$, $\boldsymbol{y}_n^{\cdot+} = \boldsymbol{y}_n(\ldots, \theta_k, \ldots, \theta_l + h, \ldots)$, $\boldsymbol{y}_n^{-\cdot} = \boldsymbol{y}_n(\ldots, \theta_k - h, \ldots, \theta_l, \ldots)$, and $\boldsymbol{y}_n^{\cdot-} = \boldsymbol{y}_n(\ldots, \theta_k, \ldots, \theta_l - h, \ldots)$. Stencil (68) requires only 2 new function evaluations compared to stencil (67)'s 4. The additional terms $\boldsymbol{y}_n$, $\boldsymbol{y}_n^{+\cdot}$, and $\boldsymbol{y}_n^{\cdot+}$ are already available from constructing the first-order sequence-space Jacobians. And the terms $\boldsymbol{y}_n^{-\cdot}$ and $\boldsymbol{y}_n^{\cdot-}$ can be computed very cheaply using the standard fake-news algorithm for first-order derivatives.

Comparison to fake-news algorithm of Auclert et al. (2021). In their seminal contribution, Auclert et al. (2021) develop a highly efficient algorithm to compute sequence-space Jacobians, showing that computing a single column of the Jacobian suffices to derive all other columns from it. For sequence-space Hessians, on the other hand, we need to evaluate one "block" of the Hessian, which requires N numerical derivatives, and is consequently substantially more expensive than computing a sequence-space Jacobian.

Why does the Hessian matrix have a higher information requirement? For Jacobians, Auclert et al. (2021) show that we only require a single piece of information to evaluate the impact of shocks on household behavior: How far in the future is the shock, i.e., what is the distance from the present to the shock. For Hessians, on the other hand, we need two pieces of information: How far in the future is the (later of the two) shock(s), and, in addition, what is the relative distance between the two shocks. We therefore cannot obtain all required information with a single numerical derivative as in the case of the Jacobian.

Nonetheless, our fake-news algorithm for sequence-space Hessians represents a substantial improvement over computing the Hessian matrices directly, which would require the evaluation of N^2 numerical derivatives.

B.5.2 Total Derivatives and General Equilibrium

Our perturbation approach to optimal stabilization policy in the dual requires the two total derivatives $\frac{d}{d\boldsymbol{\theta}}F$ and $\frac{d}{d\boldsymbol{Z}}F$. In particular, the $[k,l]$th entry of the $N \times N$ matrix $F_{\boldsymbol{\theta}}$ is given by

$$(F_{\boldsymbol{\theta}})_{[k,l]} = \sum_{n=1}^{N} e^{-\rho t_n} \frac{d^2 U_n}{d\theta_k d\theta_l} dt + (\boldsymbol{\phi}_{\text{ss}})' \frac{d^2 \boldsymbol{V}_1}{d\theta_k d\theta_l} + \vartheta_{\text{ss}} \frac{d^2 \pi_1^w}{d\theta_k d\theta_l}, \tag{69}$$

where $U_n = (u(\boldsymbol{c}_n) - v(N_n) - \frac{\delta}{2}(\pi_n^w)^2)' \boldsymbol{g}_n$. The first term in equation (69) thus captures the present discounted sum of future aggregate social welfare flows, and the second and third terms capture the timeless penalties.

So far, we have discussed how to construct the first- and second-order partial derivatives of the economic variables that comprise F. To compute total derivatives, we start with a discussion of general equilibrium.

General equilibrium. General equilibrium considerations in our model can be summarized in terms of the equilibrium map (50), which is a system of N equation, assuming for now that $X_n \in \mathbb{R}$. Given paths for policy $\boldsymbol{\theta}$ and the exogenous shock $\boldsymbol{Z}$, we can solve equation (50) for $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{\theta}, \boldsymbol{Z})$.

To compute the total derivative $F_{\boldsymbol{\theta}}$, i.e., the response in the planner's first-order condition to a perturbation in the policy path, we must take into account both the direct effect of the policy via its partial derivative and the indirect general equilibrium effects. We use the first-order derivatives $\boldsymbol{X}_{\theta_k} = -\boldsymbol{H}_{\boldsymbol{X}}^{-1}\boldsymbol{H}_{\theta_k}$. Likewise, the mixed partial derivatives are given by

$$\boldsymbol{X}_{\theta_k\theta_l} = -\boldsymbol{H}_{\boldsymbol{X}}^{-1}\boldsymbol{H}_{\theta_k\theta_l} + \boldsymbol{H}_{\boldsymbol{X}}^{-1}\boldsymbol{H}_{\boldsymbol{X}\theta_l}\boldsymbol{H}_{\boldsymbol{X}}^{-1}\boldsymbol{H}_{\theta_k}. \tag{70}$$

Total derivatives. We now summarize how total derivatives of Y_n relate to the partial derivatives we have discussed so far. For notational convenience, we drop the n subscript and instead use subscripts to denote partial derivatives. Recall that Y depends on the time paths of all aggregate inputs, $Y(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Z})$.

Lemma 29. *The total derivatives of Y are given by*

$$\frac{d^2Y}{d\theta_k d\theta_l} = \begin{pmatrix} Y_{X_1\boldsymbol{X}}\boldsymbol{X}_{\theta_l} & \ldots & Y_{X_N\boldsymbol{X}}\boldsymbol{X}_{\theta_l} \end{pmatrix} \boldsymbol{X}_{\theta_k} + Y_{\boldsymbol{X}\theta_l}\boldsymbol{X}_{\theta_k} + Y_{\boldsymbol{X}}\boldsymbol{X}_{\theta_k\theta_l} + Y_{\theta_k\boldsymbol{X}}\boldsymbol{X}_{\theta_l} + Y_{\theta_k\theta_l} \tag{71}$$

where subscripts denote partial derivatives, and likewise for the total derivatives $\frac{d^2Y}{d\theta_k dZ_l}$.

The total derivatives for $\boldsymbol{V}_1$ and π_1^w take the same form and can be computed via their second-order partial derivatives together with the general equilibrium maps, i.e., the partial derivatives of $\boldsymbol{X}$. We now have all the objects we need to implement our perturbation approach in the dual and compute optimal stabilization policy numerically.

B.5.3 Algorithm to Compute Optimal Policy in the Dual

We summarize in Algorithm 1 our fake-news algorithm to compute sequence-space Hessians and, with them, optimal stabilization policy to first order in the dual representation of the timeless Ramsey problem.

B.5.4 Accuracy and Performance

We test the accuracy of our method in a number of ways. In Appendix F.2, we show that the numerical solution of optimal policy in RANK using our perturbation method based on sequence-space Hessians is highly accurate. In RANK, we can compute optimal policy analytically. We compare this exact analytical solution to the first-order approximation of optimal policy given by $d\boldsymbol{\theta} = -F_{\boldsymbol{\theta}}^{-1}F_{\boldsymbol{Z}}d\boldsymbol{Z}$. For demand shocks, we show that the difference in optimal CPI inflation, for example, is on the order of 10^{-6}. In the case of TFP shocks, the remaining discrepancy is slightly larger, with the two optimal interest rate paths differing by about 1 basis point.

In ongoing work, we use our perturbation method to compute optimal monetary policy numerically in the analytical HANK model of Acharya et al. (2020). Since their model admits

Algorithm 1 Optimal Stabilization Policy using Sequence-Space Hessians

1: Compute stationary Ramsey plan

2: Compute sequence-space Jacobians around the stationary Ramsey plan using fake-news algorithm of Auclert et al. (2021)

3: Compute N numerical mixed partial derivatives, and ▷ use stencil (68)
   a: construct policy Hessians ▷ use Lemma 26
   b: construct distribution Hessians ▷ use Lemmas 27 and 28

4: Use Hessians to compute mixed derivatives of $\boldsymbol{H}$ and $\boldsymbol{X}$ ▷ use equations (50) and (70)

5: Compute total derivatives for $F_{\boldsymbol{\theta}}$ and $F_{\boldsymbol{Z}}$ ▷ use equations (69) and (71)

6: Compute optimal stabilization policy as $d\boldsymbol{\theta} = -F_{\boldsymbol{\theta}}^{-1} F_{\boldsymbol{Z}} d\boldsymbol{Z}$

an analytical solution for optimal policy in a heterogeneous-agent context, it represents a useful environment to benchmark the accuracy of our method.

C Labor Market Structure and Wage Phillips Curves

C.1 Wage Phillips Curve with Utility Adjustment Cost

The union's problem is associated with the Lagrangian

$$L = \int_0^\infty e^{-\rho t} \int \left[u\left(c_t\left(a,z;W_{k,t}\right) \right) - v\left(\int_0^1 \left(\frac{W_{k,t}}{W_t} \right)^{-\epsilon} N_t dk \right) - \frac{\delta}{2} \int_0^1 \left(\pi^w_{k,t} \right)^2 dk \right] g_t\left(a,z\right) d\left(a,z\right) dt$$
$$+ \int_0^\infty e^{-\rho t} \left[\mu_t \pi^w_{k,t} W_{k,t} - \rho \mu_t W_{k,t} + W_{k,t} \dot{\mu}_t \right] dt + \mu_0 W_{k,0},$$

where in the second line we already integrated by parts. Thus, the two first-order conditions are given by

$$0 = \int u'(c_t) \frac{\partial c_t(a,z;W_{k,t})}{\partial W_{k,t}} g_t\left(a,z\right) d\left(a,z\right) + \epsilon v'(N_t) \frac{N_t}{W_t} + \mu_t \pi^w_{k,t} - \rho \mu_t + \dot{\mu}_t$$
$$0 = -\delta \pi^w_{k,t} + \mu_t W_{k,t},$$

as well as the initial condition $\mu_0 = 0$. By the envelope theorem, we have

$$\frac{\partial c_t(a,z;W_{k,t})}{\partial W_{k,t}} = \frac{1}{P_t}(1+\tau^L)(1-\epsilon) z_t N_t.$$

Defining

$$\Lambda_t = \int z u'(c_t\left(a,z\right)) g_t\left(a,z\right) d\left(a,z\right),$$

the first FOC becomes

$$0 = (1+\tau^L)(1-\epsilon) w_t N_t \Lambda_t + \epsilon v'(N_t) N_t + \mu_t \dot{W}_t - \rho W_t \mu_t + W_t \dot{\mu}_t.$$

Differentiating the second FOC yields

$$\mu_t \dot{W}_t + W_t \dot{\mu}_t = \delta \dot{\pi}^w_t.$$

Plugging back into the first FOC, we arrive at

$$0 = (1+\tau^L)(1-\epsilon) w_t N_t \Lambda_t + \epsilon v'(N_t) N_t - \rho \delta \pi^w_t + \delta \dot{\pi}^w_t$$

which yields the result after rearranging.

D RANK with Wage Rigidity

In this Appendix, we present a self-contained treatment of optimal monetary policy in RANK. We leverage our timeless Ramsey approach to give an exact, non-linear characterization, which we leverage in Section 4 of the main text to compare optimal policy in HANK to the RANK benchmark.

D.1 Model

The representative household has preferences over consumption and labor. We assume that the household's labor decision is intermediated by a continuum of $k \in [0,1]$ unions, which we further describe below. Preferences are thus given by

$$\int_0^\infty e^{-\rho t} \Big[u(C_t) - V(N_{k,t}, \pi_{k,t}^w) \Big] dt,$$

where C_t denotes consumption of the final consumption good and $V(n_{k,t}, \pi_{k,t}^w)$ denotes disutility from work. The representative household's budget constraint is given by

$$\dot{A}_t = r_t A_t + E_t + \tau_t - C_t,$$

where A_t is the aggregate bond position, r_t the real interest rate, and E_t denotes real labor income. Finally, τ_t denotes a real lump-sum rebate to households.

Labor market structure. The labor market structure and union problem in the RANK benchmark are unchanged relative to HANK except for one important difference: Union k maximizes stakeholder value, which in HANK is an appropriately weighted average of its members' utility flows. In RANK, on the other hand, there is a single representative household. Consequently, the marginal cost term in the New Keynesian wage Phillips curve that obtains in RANK simply features the marginal utility of consumption of the representative household, i.e.,

$$\dot{\pi}_t^w = \rho \pi_t^w + \frac{\epsilon}{\delta} \Big[(1+\tau^L) \frac{\epsilon - 1}{\epsilon} w_t u'(C_t) - v'(N_t) \Big] N_t.$$

instead of the weighted average of marginal utilities, Λ_t, which features in the Phillips curve in HANK.

Equilibrium. The characterization of firms in RANK is identical to that in HANK. As are the details of fiscal and monetary policy. To formally define equilibrium and RANK and state the implementability conditions for the Ramsey problem, we first characterize the representative household's optimal consumption behavior in terms of an aggregate consumption Euler equation. Under isoelastic preferences, which we assume in this Appendix, we simply arrive at the standard differential equation

$$\frac{\dot{C}_t}{C_t} = \frac{r_t - \rho}{\gamma}.$$

The equilibrium conditions of the RANK economy are then given by

$$\begin{aligned}
\dot{C}_t &= \frac{r_t - \rho}{\gamma} C_t \\
Y_t &= A_t N_t \\
Y_t &= C_t \\
B_t &= 0 \\
r_t &= i_t - \pi_t \\
w_t &= A_t \\
\pi_t &= \pi_t^w - \frac{\dot{A}_t}{A_t} \\
\dot{\pi}_t^w &= \rho \pi_t^w + \frac{\epsilon}{\delta}\left[(1+\tau^L)\frac{\epsilon - 1}{\epsilon} w_t u'(C_t) - v'(N_t)\right] N_t
\end{aligned}$$

We drop the bond market clearing condition. We substitute in for the real wage as well as CPI inflation. And we drop output. We summarize the resulting implementability conditions in the following Lemma.

Lemma 30. *The implementability conditions that a Ramsey planner faces in RANK can be summarized as*

$$\begin{aligned}
\dot{C}_t &= \frac{r_t - \rho}{\gamma} C_t \\
C_t &= A_t N_t \\
r_t &= i_t - \pi_t^w + \frac{\dot{A}_t}{A_t} \\
\dot{\pi}_t^w &= \rho \pi_t^w + \frac{\epsilon}{\delta}\left[(1+\tau^L)\frac{\epsilon - 1}{\epsilon} A_t u'(C_t) - v'(N_t)\right] N_t
\end{aligned}$$

It is important to note that the implementability conditions of Lemma 30 could be simplified further. We refrain from doing so in order to maintain as parallel a structure to the Ramsey problem in HANK as possible.

D.2 Optimal Policy

We associate the standard Ramsey problem in the primal with the following Lagrangian, where we drop time subscripts for convenience,

$$
\begin{aligned}
L = \int_0^\infty e^{-\rho t} \Bigg\{ & \frac{1}{1-\gamma} C^{1-\gamma} - v(N) - \frac{\delta}{2} (\pi^w)^2 \\
& + \phi \left[\frac{1}{\gamma} \left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) C - \dot{C} \right] \\
& + \mu \Big[AN - C \Big] \\
& + \vartheta \left[\rho \pi^w + \frac{\epsilon}{\delta} \left((1+\tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C) - v'(N) \right) N - \dot{\pi}^w \right] \Bigg\} dt
\end{aligned}
$$

Crucially, both C_0 and π_0^w are *free* from the planner's perspective.

Lemma 31. *The first-order conditions for optimal monetary policy in RANK are given by*

$$
0 = C^{-\gamma} - \mu + \phi \frac{1}{\gamma} \left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) - \rho \phi + \dot{\phi} + \vartheta \frac{\epsilon}{\delta} (1+\tau^L) \frac{\epsilon - 1}{\epsilon} A u''(C) N \tag{72}
$$

$$
0 = -\delta \pi^w - \phi \frac{1}{\gamma} C + \vartheta \rho - \rho \vartheta + \dot{\vartheta} \tag{73}
$$

$$
0 = -v'(N) + \mu A + \vartheta \frac{\epsilon}{\delta} \left((1+\tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C) - v'(N) \right) - \vartheta \frac{\epsilon}{\delta} v''(N) N \tag{74}
$$

$$
0 = \phi \frac{1}{\gamma} C, \tag{75}
$$

with initial conditions

$$
\begin{aligned}
0 &= \phi_0 \\
0 &= \vartheta_0.
\end{aligned}
$$

We see that we must have $\phi = 0$ for all t. This allows us to simplify the first-order conditions and arrive at

$$
\begin{aligned}
0 &= C^{-\gamma} - \mu + \vartheta \frac{\epsilon}{\delta} (1+\tau^L) \frac{\epsilon - 1}{\epsilon} A u''(C) N \\
\dot{\vartheta} &= -\delta \pi^w \\
0 &= -v'(N) + \mu A + \vartheta \frac{\epsilon}{\delta} (1+\tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C) - \vartheta \frac{\epsilon}{\delta} v'(N) - \vartheta \frac{\epsilon}{\delta} v''(N) N
\end{aligned}
$$

with initial condition $\vartheta_0 = 0$.

Proof. We now integrate by parts and consider a general functional perturbation, yielding

$$
\begin{aligned}
L = \int_0^\infty e^{-\rho t} \Bigg\{ & \frac{1}{1-\gamma}(C + \alpha h_C)^{1-\gamma} - v(N + \alpha h_N) - \frac{\delta}{2}(\pi^w + \alpha h_\pi)^2 \\
& + \phi \left[\frac{1}{\gamma} \left(i + \alpha h_i - \pi^w - \alpha h_\pi + \frac{\dot{A}}{A} - \rho \right) (C + \alpha h_C) \right] \\
& + \mu \left[A(N + \alpha h_N) - C - \alpha h_C \right] \\
& + \vartheta \left[\rho(\pi^w + \alpha h_\pi) + \frac{\epsilon}{\delta} \left((1 + \tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C + \alpha h_C) - v'(N + \alpha h_N) \right) (N + \alpha h_N) \right] \\
& - \rho \phi (C + \alpha h_C) + (C + \alpha h_C) \dot{\phi} - \rho \vartheta (\pi^w + \alpha h_\pi) + (\pi^w + \alpha h_\pi) \dot{\vartheta} \Bigg\} dt + \phi_0 (C_0 + \alpha h_{C,0}) + \vartheta_0 (\pi_0^w + \alpha h_\pi
\end{aligned}
$$

Working out the Gateaux derivatives and employing the fundamental lemma of the calculus of variations, we arrive at the following

$$
\begin{aligned}
0 = \int_0^\infty e^{-\rho t} \Bigg\{ & C^{-\gamma} h_C - v'(N) h_N - \delta \pi^w h_\pi \\
& + \phi \left[\frac{1}{\gamma} \left(h_i - h_\pi \right) C + \frac{1}{\gamma} \left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) h_C \right] \\
& + \mu \left[A h_N - h_C \right] \\
& + \vartheta \Bigg[\rho h_\pi + \frac{\epsilon}{\delta} \left((1 + \tau^L) \frac{\epsilon - 1}{\epsilon} A u''(C) h_C - v''(N) h_N \right) N \\
& \qquad + \frac{\epsilon}{\delta} \left((1 + \tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C) - v'(N) \right) h_N \Bigg] \\
& - \rho \phi h_C + h_C \dot{\phi} - \rho \vartheta h_\pi + h_\pi \dot{\vartheta} \Bigg\} dt + \phi_0 h_{C,0} + \vartheta_0 h_{\pi,0}
\end{aligned}
$$

Grouping terms,

$$
\begin{aligned}
0 = \int_0^\infty e^{-\rho t} \Bigg\{ & C^{-\gamma} h_C - \mu h_C + \phi \frac{1}{\gamma}\left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) h_C - \rho \phi h_C + h_C \dot{\phi} + \vartheta \frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon - 1}{\epsilon} A u''(C) N h_C \\
& - \delta \pi^w h_\pi - \phi \frac{1}{\gamma} C h_\pi + \vartheta \rho h_\pi - \rho \vartheta h_\pi + h_\pi \dot{\vartheta} \\
& - v'(N) h_N + \mu A h_N + \vartheta \frac{\epsilon}{\delta}\left((1+\tau^L)\frac{\epsilon - 1}{\epsilon} A u'(C) - v'(N) \right) h_N - \vartheta \frac{\epsilon}{\delta} v''(N) h_N N \\
& + \phi \frac{1}{\gamma} C h_i \Bigg\} dt + \phi_0 h_{C,0} + \vartheta_0 h_{\pi,0}
\end{aligned}
$$

The fundamental lemma of the calculus of variations yields the desired result.

■

Lemma 32. *The stationary Ramsey plan in RANK satisfies*

$$
\begin{aligned}
\pi^w_{ss} &= 0 \\
i_{ss} &= \rho \\
N_{ss} &= \left[(1+\tau^L)\frac{\epsilon - 1}{\epsilon} \right]^{\frac{1}{\gamma+\eta}} \\
C_{ss} &= N_{ss} \\
\vartheta_{ss} &= \frac{1 - (1+\tau^L)\frac{\epsilon-1}{\epsilon}}{(\gamma+\eta)\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}} \\
\mu_{ss} &= \frac{\gamma}{\gamma+\eta}(N_{ss})^\eta + \frac{\eta}{\gamma+\eta}(C_{ss})^{-\gamma}.
\end{aligned}
$$

Importantly, we see that $\vartheta_{\text{ss}} = 0$ if and only if an appropriate employment subsidy is in place, so that $(1+\tau^L)\frac{\epsilon-1}{\epsilon} = 1$.

D.3 Timeless Dual Lagrangian

In the following, we leverage our timeless Ramsey approach to give a novel, non-linear characterization of optimal monetary policy in RANK. We associate the timeless Ramsey problem in the dual with the Lagrangian

$$
L^{\text{TD}}(\vartheta) = \int_0^\infty e^{-\rho t} \left\{ \frac{1}{1-\gamma} C_t^{1-\gamma} - v(N_t) - \frac{\delta}{2}(\pi^w_t)^2 \right\} dt \underbrace{-\vartheta \pi^w_0}_{\text{Inflation Target}}
$$

Lemma 33. *The timeless dual Ramsey problem in RANK is time consistent. In the absence of shocks, the Ramsey planner has no incentive to deviate from the stationary Ramsey plan. That is,*

$$\frac{d}{d\boldsymbol{\theta}} L^{TD}(\vartheta)\bigg|_{ss} = 0.$$

Proof. Suppose we differentiate

$$\frac{d}{d\boldsymbol{\theta}} L^{\mathrm{TD}}(\vartheta) = \int_0^\infty e^{-\rho t} \left\{ C_t^{-\gamma} \frac{d}{d\boldsymbol{\theta}} - N_t^{\eta} \frac{dN_t}{d\boldsymbol{\theta}} - \delta \pi_t^w \frac{d\pi_t^w}{d\boldsymbol{\theta}} \right\} dt - \vartheta \frac{d\pi_0^w}{d\boldsymbol{\theta}}$$

Next, we evaluate at the stationary Ramsey plan. This yields

$$\begin{aligned} \frac{d}{d\boldsymbol{\theta}} L^{\mathrm{TD}}(\vartheta) &= \int_0^\infty e^{-\rho t} \left\{ C^{-\gamma} \frac{dC_t}{d\boldsymbol{\theta}} - N^{\eta} \frac{dN_t}{d\boldsymbol{\theta}} \right\} dt - \vartheta \frac{d\pi_0^w}{d\boldsymbol{\theta}} \\ &= \int_0^\infty e^{-\rho t} \left\{ C^{-\gamma} \left[\frac{dC_t}{d\boldsymbol{\theta}} - (1+\tau^L) \frac{\epsilon - 1}{\epsilon} \frac{dN_t}{d\boldsymbol{\theta}} \right] \right\} dt - \vartheta \frac{d\pi_0^w}{d\boldsymbol{\theta}} \end{aligned}$$

Next, from $C_t = A_t N_t$, we have when evaluated at the stationary Ramsey plan that

$$\frac{dC_t}{d\boldsymbol{\theta}} = A \frac{dN_t}{d\boldsymbol{\theta}}.$$

Thus,

$$\begin{aligned} \frac{d}{d\boldsymbol{\theta}} L^{\mathrm{TD}}(\vartheta) &= \int_0^\infty e^{-\rho t} \left\{ C^{-\gamma} \frac{dC_t}{d\boldsymbol{\theta}} - N^{\eta} \frac{dN_t}{d\boldsymbol{\theta}} \right\} dt - \vartheta \frac{d\pi_0^w}{d\boldsymbol{\theta}} \\ &= C^{-\gamma} \left[1 - (1+\tau^L) \frac{\epsilon - 1}{\epsilon} \right] \int_0^\infty e^{-\rho t} \frac{dC_t}{d\boldsymbol{\theta}} dt - \vartheta \frac{d\pi_0^w}{d\boldsymbol{\theta}}. \end{aligned}$$

Next, we use the Phillips curve. With $\lim_{T\to\infty} \pi_T^w = 0$, we have in integral form

$$\dot{\pi}_t^w = \rho \pi_t^w + \frac{\epsilon}{\delta} \left[(1+\tau^L) \frac{\epsilon - 1}{\epsilon} A_t u'(C_t) - v'(N_t) \right] N_t$$

$$\pi_t^w = -\int_t^\infty e^{-\rho(s-t)} \frac{\epsilon}{\delta} \left[(1+\tau^L) \frac{\epsilon - 1}{\epsilon} A_s u'(C_s) - v'(N_s) \right] N_s$$

Thus, we have

$$\frac{d\pi_0^w}{d\boldsymbol{\theta}} = -\int_0^\infty e^{-\rho(s-0)}\frac{\epsilon}{\delta}\left[(1+\tau^L)\frac{\epsilon-1}{\epsilon}(1-\gamma)C^{-\gamma}\frac{dC_s}{d\boldsymbol{\theta}} - (1+\eta)N^\eta\frac{dN_s}{d\boldsymbol{\theta}}\right]ds$$
$$= -\frac{\epsilon}{\delta}\left[(1+\tau^L)\frac{\epsilon-1}{\epsilon}(1-\gamma)C^{-\gamma} - (1+\eta)N^\eta\right]\int_0^\infty e^{-\rho t}\frac{dC_t}{d\boldsymbol{\theta}}dt$$
$$= \frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}C^{-\gamma}(\gamma+\eta)\int_0^\infty e^{-\rho t}\frac{dC_t}{d\boldsymbol{\theta}}dt$$

Thus, we have

$$\frac{d}{d\boldsymbol{\theta}}L^{\text{TD}}(\vartheta) = \overbrace{C^{-\gamma}\left[1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}\right]\int_0^\infty e^{-\rho t}\frac{dC_t}{d\boldsymbol{\theta}}dt}^{\text{Marginal benefit from time-inconsistent deviations}}$$
$$\underbrace{-\vartheta\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}C^{-\gamma}(\gamma+\eta)\int_0^\infty e^{-\rho t}\frac{dC_t}{d\boldsymbol{\theta}}dt}_{\text{Marginal cost of time-inconsistent deviations under timeless penalty / inflation target}}$$

Finally, we now have

$$0 = C^{-\gamma}\left[1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}\right] - \vartheta\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}C^{-\gamma}(\gamma+\eta)$$
$$= \left[1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}\right] - \frac{1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}}{(\gamma+\eta)\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}}\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}(\gamma+\eta)$$
$$= \left[1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}\right] - \left[1-(1+\tau^L)\frac{\epsilon-1}{\epsilon}\right],$$

which concludes the proof.

■

Remark 34. Our constructive proof of Lemma 33 characterizes clearly the *marginal benefit* from time-inconsistent deviations from the stationary Ramsey plan. And it also shows clearly how the timeless penalty, the *marginal cost* of deviations, exactly offsets the marginal benefit. Importantly, we see here in closed-form what the economic determinants are of the marginal benefit and the timeless penalty.

D.4 Retracing Classical RANK Results

We are now ready to use our apparatus to retrace the classical analysis of optimal monetary stabilization policy in RANK. In this subsection, we restate several of the classical results in an exact, non-linear form.

To that end, we now consider a version of the baseline RANK model with demand, productivity, and cost-push shocks. Formally, we assume the following exogenous processes

$$\dot{A}_t = \xi_A(A - A_t)$$

$$\dot{\epsilon}_t = \xi_\epsilon(\epsilon - \epsilon_t)$$

$$\dot{\rho}_t = \xi_\rho(\rho - \rho_t)$$

where A, ϵ, and ρ are the steady-state constant levels. We will consider one-time unanticipate ("MIT") shocks at time $t = 0$, so that A_0, ϵ_0, and ρ_0 jump and subsequently revert back to their steady-state levels following the above laws of motion. With this enriched structure, the implementability conditions for the optimal policy Ramsey problem become

$$\dot{C}_t = \frac{1}{\gamma}\left(i_t - \pi_t^w + \frac{\dot{A}_t}{A_t} - \rho_t\right)C_t$$

$$C_t = A_t N_t - G_t$$

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta}\left[(1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}A_t u'(C_t) - v'(N_t)\right]N_t$$

In much of the standard RANK literature, e.g., Clarida et al. (1999), optimal policy analysis drops the IS equation as an implementability condition and then proceeds to derive *targeting rules* for inflation and output (gaps). In the following, our goal is to retrace this classical analysis in our setting. We will leverage the results we derive below in Section 4 of the main text to compare optimal policy and targeting rules across RANK and HANK.

We drop the IS equation and use the resource constraint to solve out for N_t. Following Galí (2015), we define the *natural level of output*, denoted Y_t^n, as the equilibrium level of output under flexible prices. From the Phillips curve, which is in our setting given by

$$Y_t^n = \left[(1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}A_t^{1+\eta}\right]^{\frac{1}{\gamma+\eta}}. \tag{76}$$

Going back to the Phillips curve and using the resource constraint with $Y_t = A_t N_t = C_t$, we have

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta}\left[\left(Y_t^n\right)^{\gamma+\eta} - Y_t^{\gamma+\eta}\right]Y_t^{1-\gamma}A_t^{-1-\eta} \tag{77}$$

which is our sole remaining implementability condition and features all three shocks: A_t, ϵ_t, and ρ_t.

The planner's Ramsey problem can now be associated with the Lagrangian

$$L = \int_0^\infty e^{-\int_0^t \rho_s ds} \Bigg\{ u(Y_t) - v\left(\frac{Y_t}{A_t}\right) - \frac{\delta}{2}(\pi_t^w)^2 \\ + \vartheta_t \left[\rho_t \pi^w + \frac{\epsilon_t}{\delta}\left((1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t} A_t u'(Y_t) - v'\left(\frac{Y_t}{A_t}\right)\right)\frac{Y_t}{A_t} - \dot{\pi}_t^w \right] \Bigg\} dt$$

We now state the main result of this appendix: a non-linear targeting rule for optimal monetary policy in RANK under demand, TFP, and cost-push shocks.

Proposition 35. (Optimal Policy Targeting Rules / Divine Coincidence in RANK)
a) (Targeting Rule) Optimal monetary policy in RANK is fully characterized by the non-linear targeting rule

$$Y_t = Y_t^n \left(\frac{\frac{1}{1+\tau^L}\frac{\epsilon_t}{\epsilon_t - 1} + \frac{\epsilon_t}{\delta}\vartheta_t(1-\gamma)}{1 + \frac{\epsilon_t}{\delta}\vartheta_t(1+\eta)} \right)^{\frac{1}{\gamma+\eta}} \tag{78}$$

b) (Divine Coincidence) Suppose there are no cost-push shocks, i.e., $\epsilon_t = \epsilon$, and we implement an employment subsidy so that $(1+\tau^L)\frac{\epsilon-1}{\epsilon} = 1$. We have

$$Y_t = Y_t^n \left(\frac{1 + \frac{\epsilon}{\delta}\vartheta_t(1-\gamma)}{1 + \frac{\epsilon}{\delta}\vartheta_t(1+\eta)} \right)^{\frac{1}{\gamma+\eta}}. \tag{79}$$

A solution to the non-linear Ramsey plan is then given by $Y_t = Y_t^n$, $\vartheta_t = 0$, and $\pi_t^w = 0$.

Proof. Crucially, both Y_0 and π_0^w are *free* from the planner's perspective. We start by integrating by parts, yielding

$$L = \int_0^\infty e^{-\int_0^t \rho_s ds} \Bigg\{ u(Y_t) - v\left(\frac{Y_t}{A_t}\right) - \frac{\delta}{2}(\pi_t^w)^2 \\ + \vartheta_t \left[\rho_t \pi^w + \frac{\epsilon_t}{\delta}\left((1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t} A_t u'(Y_t) - v'\left(\frac{Y_t}{A_t}\right)\right)\frac{Y_t}{A_t} \right] \\ - \rho_t \vartheta_t \pi_t^w + \pi_t^w \dot{\vartheta}_t \Bigg] \Bigg\} dt + \vartheta_0 \pi_0^w$$

The two first-order conditions are then given by

$$\begin{aligned}0 =& u'(Y_t) - v'\left(\frac{Y_t}{A_t}\right)\frac{1}{A_t} + \frac{\epsilon_t}{\delta}\vartheta_t\left[(1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}A_t u''(Y_t) - v''\left(\frac{Y_t}{A_t}\right)\frac{1}{A_t}\right]\frac{Y_t}{A_t} \\ &+ \frac{\epsilon_t}{\delta}\vartheta_t\left((1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}A_t u'(Y_t) - v'\left(\frac{Y_t}{A_t}\right)\right)\frac{1}{A_t}\end{aligned}$$

for output and $\dot{\vartheta}_t = \delta\pi_t^w$ for the multiplier.

We now simplify the first condition, which will take the form of a targeting rule, as discussed in much of the classical optimal policy analysis in RANK. With isoelastic preferences, we have

$$0 = Y_t^{-\gamma} - Y_t^{\eta}A_t^{-\eta-1} + \frac{\epsilon_t}{\delta}\vartheta_t\left((1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}(1-\gamma)Y_t^{-\gamma} - (1+\eta)Y_t^{\eta}A_t^{-\eta-1}\right)$$

Further rearranging yields

$$0 = A_t^{1+\eta} - Y_t^{\gamma+\eta} + \frac{\epsilon_t}{\delta}\vartheta_t(1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}(1-\gamma)A_t^{1+\eta} - \frac{\epsilon_t}{\delta}\vartheta_t(1+\eta)Y_t^{\gamma+\eta}$$

or simply

$$\left[1 + \frac{\epsilon_t}{\delta}\vartheta_t(1+\eta)\right]^{\frac{1}{\gamma+\eta}} Y_t = \left[\left(\frac{1}{1+\tau^L}\frac{\epsilon_t}{\epsilon_t - 1} + \frac{\epsilon_t}{\delta}\vartheta_t(1-\gamma)\right)(1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t}A_t^{1+\eta}\right]^{\frac{1}{\gamma+\eta}}$$

Using the definition of natural output, we therefore have

$$\left[1 + \frac{\epsilon_t}{\delta}\vartheta_t(1+\eta)\right]^{\frac{1}{\gamma+\eta}} Y_t = \left(\frac{1}{1+\tau^L}\frac{\epsilon_t}{\epsilon_t - 1} + \frac{\epsilon_t}{\delta}\vartheta_t(1-\gamma)\right)^{\frac{1}{\gamma+\eta}} Y_t^n$$

or simply

■

Importantly, the targeting rule of Proposition 35 echoes the seminal result of the standard New Keynesian framework, that Divine Coincidence obtains unless there are cost-push shocks. In the presence of only productivity and demand shocks, the planner perceives no tradeoff between inflation and output.

E Optimal Policy under Discretion

The optimal planning problem under discretion at time s is formally associated with the Lagrangian

$$
\begin{aligned}
L^{\mathrm{D}}(\boldsymbol{g}_s) = \min_{\boldsymbol{\phi}_s,\boldsymbol{\chi}_s,\boldsymbol{\lambda}_s,\mu_s,\vartheta_s} \max_{\boldsymbol{V}_s,\boldsymbol{c}_{s,[2:J]},\boldsymbol{g}_{s+1},\pi^w_s,N_s,i_s} \sum_{n=s}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ \\
&+ u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_t - v(N_n)\mathbf{1}'\boldsymbol{g}_t - \frac{\delta}{2}(\pi^w_n)^2\mathbf{1}'\boldsymbol{g}_t \\
&+ \boldsymbol{\phi}'_n\left[-\rho \boldsymbol{V}_n + \frac{\boldsymbol{V}_{n+1}-\boldsymbol{V}_n}{dt} + u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} - v(N_n) - \frac{\delta}{2}(\pi^w_n)^2 \right] \\
&+ \boldsymbol{\phi}'_n \boldsymbol{A}^z \boldsymbol{V}_n + \sum_{i\geq 2} \phi_{i,n}\left(i_n a_i - \pi^w_n a_i + \frac{A_{n+1}-A_n}{dtA_n} a_i + z_i A_n N_n - c_{n,i}\right) \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n \\
&+ \boldsymbol{\chi}'_{n,[2:J]}\left[u'(\boldsymbol{c}_{n,[2:J]}) - \left(\frac{\boldsymbol{D}_a}{da}\boldsymbol{V}_{n+1}\right)_{[2:J]} \right] \\
&- \boldsymbol{\lambda}'_n \frac{\boldsymbol{g}_{n+1}-\boldsymbol{g}_n}{dt} + \boldsymbol{\lambda}'_n (\boldsymbol{A}^z)'\boldsymbol{g}_n \\
&+ \sum_{i\geq 2} \lambda_{n,i} \frac{\boldsymbol{D}'_{a,[i,:]}}{da}\left[\begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi^w_n \boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n}\boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \cdot \boldsymbol{g}_n \right] \Bigg\} d\boldsymbol{x} \\
&+ \mu_n \left[\boldsymbol{c}'_n \boldsymbol{g}_n d\boldsymbol{x} - A_n N_n \right] \\
&+ \vartheta_n \left[-\frac{\pi^w_{n+1}-\pi^w_n}{dt} + \rho\pi^w_n + \frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L) A_n (\boldsymbol{z}\cdot u'(\boldsymbol{c}_n))'\boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) N_n \right] \Bigg\} dt
\end{aligned}
$$

where the superscript D denotes the planning problem under discretion. The planner takes as given an initial condition for the cross-sectional distribution, $\boldsymbol{g}_s$.

Unlike in the Ramsey problem with commitment, we only integrate by parts the *state variables* of the problem, and not those terms associated with forward-looking constraints. That is, we use

$$
\begin{aligned}
-\sum_{n=s}^{N-1} e^{-\rho t_n} \boldsymbol{\lambda}'_n \frac{\boldsymbol{g}_{n+1}-\boldsymbol{g}_n}{dt} &= \sum_{n=s}^{N-1} e^{-\rho t_n} \frac{\boldsymbol{\lambda}'_n - e^{\rho dt}\boldsymbol{\lambda}'_{n-1}}{dt} \boldsymbol{g}_n \\
&+ \frac{1}{dt} e^{\rho dt} \boldsymbol{\lambda}'_{s-1} \boldsymbol{g}_s - \frac{1}{dt} e^{\rho dt} e^{-\rho t_N} \boldsymbol{\lambda}'_{N-1} \boldsymbol{g}_N
\end{aligned}
$$

and rewrite the Lagrangian as

$$
\begin{aligned}
L^{\mathrm{D}}(\boldsymbol{g}_s) = & \min_{\boldsymbol{\phi}_s,\boldsymbol{\chi}_s,\boldsymbol{\lambda}_s,\mu_s,\vartheta_s} \max_{\boldsymbol{V}_s,\boldsymbol{c}_{s,[2:J]},\boldsymbol{g}_{s+1},\pi^w_s,N_s,i_s} \sum_{n=s}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ \\
& + u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix}' \boldsymbol{g}_t - v(N_n)\mathbf{1}'\boldsymbol{g}_t - \frac{\delta}{2}(\pi^w_n)^2 \mathbf{1}'\boldsymbol{g}_t \\
& + \boldsymbol{\phi}_n' \left[-\rho \boldsymbol{V}_n + \frac{\boldsymbol{V}_{n+1}-\boldsymbol{V}_n}{dt} + u\begin{pmatrix} i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1}-A_n}{dtA_n} a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]} \end{pmatrix} - v(N_n) - \frac{\delta}{2}(\pi^w_n)^2 \right] \\
& + \boldsymbol{\phi}_n' \boldsymbol{A}^z \boldsymbol{V}_n + \sum_{i\geq 2} \phi_{i,n} \left(i_n a_i - \pi^w_n a_i + \frac{A_{n+1}-A_n}{dtA_n} a_i + z_i A_n N_n - c_{n,i} \right) \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n \\
& + \boldsymbol{\chi}_{n,[2:J]}' \left[u'(\boldsymbol{c}_{n,[2:J]}) - \left(\frac{\boldsymbol{D}_a}{da} \boldsymbol{V}_{n+1} \right)_{[2:J]} \right] \\
& + e^{-\rho t_n} \frac{\boldsymbol{\lambda}_n' - e^{\rho dt} \boldsymbol{\lambda}_{n-1}'}{dt} \boldsymbol{g}_n + \boldsymbol{\lambda}_n' (\boldsymbol{A}^z)' \boldsymbol{g}_n \\
& + \sum_{i\geq 2} \lambda_{n,i} \frac{\boldsymbol{D}_{a,[i,:]}'}{da} \left[\begin{pmatrix} 0 \\ i_n \boldsymbol{a}_{[2:J]} - \pi^w_n \boldsymbol{a}_{[2:J]} + \frac{A_{n+1}-A_n}{dtA_n} \boldsymbol{a}_{[2:J]} + \boldsymbol{z}_{[2:J]} A_n N_n - \boldsymbol{c}_{n,[2:J]} \end{pmatrix} \cdot \boldsymbol{g}_n \right] \Bigg\} d\boldsymbol{x} \\
& + \mu_n \left[\boldsymbol{c}_n' \boldsymbol{g}_n d\boldsymbol{x} - A_n N_n \right] \\
& + \vartheta_n \left[-\frac{\pi^w_{n+1} - \pi^w_n}{dt} + \rho \pi^w_n + \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1+\tau^L) A_n (\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) N_n \right] \Bigg\} dt \\
& + e^{\rho dt} \boldsymbol{\lambda}_{s-1}' \boldsymbol{g}_s d\boldsymbol{x} - e^{\rho dt} e^{-\rho t_N} \boldsymbol{\lambda}_{N-1}' \boldsymbol{g}_N d\boldsymbol{x}
\end{aligned}
$$

We now characterize the first-order optimality conditions associated with the planning problem under discretion.

Derivative for $\boldsymbol{V}_s$. We have

$$
0 = -\rho \boldsymbol{\phi}_s - \frac{1}{dt} \boldsymbol{\phi}_s + (\boldsymbol{A}^z)' \boldsymbol{\phi}_s + \frac{1}{da} (\boldsymbol{\phi}_s \boldsymbol{D}_a)' \boldsymbol{s}_s - e^{\rho dt} \frac{\boldsymbol{D}_a'}{da} \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{s-1,[2:J]} \end{pmatrix}
$$

or simply

$$
0 = -\rho \boldsymbol{\phi}_s - \frac{1}{dt} \boldsymbol{\phi}_s + \boldsymbol{A}' \boldsymbol{\phi}_s - e^{\rho dt} \frac{\boldsymbol{D}_a'}{da} \begin{pmatrix} 0 \\ \boldsymbol{\chi}_{s-1,[2:J]} \end{pmatrix}
$$

Consider the last term in this equation. The household's consumption FOC says that consumption today is a function of "expected" future value, which therefore uses $\boldsymbol{V}_{n+1}$. The planner under discretion takes the future value $\boldsymbol{V}_{n+1}$ as given. And the planner is constrained by the competitive equilibrium condition that households make consumption decisions *purely* in terms of $\boldsymbol{V}_{n+1}$. By the household's first-order condition, then, $\boldsymbol{c}_n$ is pinned down as a function of $\boldsymbol{V}_{n+1}$.

We now see from this that, in the continuous-time limit with $dt \to 0$, we must have

$$\boldsymbol{\phi}_s = 0.$$

This is the proper boundary condition for the formal continuous-time problem under discretion. Moreover, from the consumption FOC in the Lagrangian, we also have

$$0 = \frac{\boldsymbol{D}_a'}{da}\begin{pmatrix} 0 \\ \boldsymbol{\chi}_{s-1,[2:J]} \end{pmatrix}$$

for all s.

Derivative for $\boldsymbol{g}_{s+1}$. We have

$$\begin{aligned} 0 =& u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}'_{s+1} - e^{\rho dt}\boldsymbol{\lambda}'_s}{dt} + (\boldsymbol{\lambda}'_{s+1}(\boldsymbol{A}^z)')' \\ &+ \frac{d}{d\boldsymbol{g}_{s+1}}\left[\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\lambda}_{s+1})'[\boldsymbol{s}_{s+1}\cdot\boldsymbol{g}_{s+1}]\right] + \vartheta_{s+1}N_{s+1}\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_{s+1}\boldsymbol{z}\cdot u'(\boldsymbol{c}_{s+1}) \end{aligned}$$

Now we work out the remaining derivative,

$$\frac{d}{d\boldsymbol{g}_{s+1}}\left[\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\lambda}_{s+1})'[\boldsymbol{s}_{s+1}\cdot\boldsymbol{g}_{s+1}]\right] = \boldsymbol{A}^a\boldsymbol{\lambda}_{s+1}.$$

Thus, we have

$$\begin{aligned} 0 =& u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}'_{s+1} - e^{\rho dt}\boldsymbol{\lambda}'_s}{dt} \\ &+ \boldsymbol{A}\boldsymbol{\lambda}_{s+1} + \vartheta_{s+1}N_{s+1}\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_{s+1}\boldsymbol{z}\cdot u'(\boldsymbol{c}_{s+1}) \end{aligned}$$

Derivative $\boldsymbol{c}_{s,[2:J]}$. Notice that the planner under commitment also just solves a static problem for consumption at every time step. In other words, the choice of consumption today doesn't "bind" the planner tomorrow in any way under commitment. Therefore, we again have

$$\begin{aligned} 0 =& u'(c_{s,i})g_{s,i} + \mu_s g_{s,i} + u'(c_{s,i})\phi_{s,i} + u''(c_{s,i})\chi_{s,i} + \vartheta_s N_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s z_i u''(c_{s,i})g_{s,i} \\ &- \frac{1}{da}\phi_{s,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}_s - \frac{1}{da}g_{s,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{\lambda}_s \end{aligned}$$

Derivative π_n^w. We have

$$0 = \Big[-u'(c_{s,1})g_{s,1}a_1 - \mu_s g_{s,1}a_1 - \delta\pi_s^w \mathbf{1}'\boldsymbol{g}_s - \phi_{s,1}u'(c_{s,1})a_1 - \delta\pi_s^w \boldsymbol{\phi}_s'\mathbf{1}\Big]d\boldsymbol{x}$$

$$-\vartheta_s \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}a_1 d\boldsymbol{x}$$

$$+\Big[-\sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}\lambda_{s,i}\frac{\boldsymbol{D}'_{a,[i,:]}}{da}\Big[\begin{pmatrix}0\\ -\boldsymbol{a}_{[2:J]}\end{pmatrix}\cdot \boldsymbol{g}_s\Big]d\boldsymbol{x}$$

$$+\frac{1}{dt}\vartheta_s$$

Thus, we have

$$0 = \Big[-u'(c_{s,1})g_{s,1}a_1 - \mu_s g_{s,1}a_1 - \delta\pi_s^w \mathbf{1}'\boldsymbol{g}_s - \phi_{s,1}u'(c_{s,1})a_1 - \delta\pi_s^w \boldsymbol{\phi}_s'\mathbf{1}\Big]d\boldsymbol{x}$$

$$-\vartheta_s \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}a_1 d\boldsymbol{x}$$

$$-\sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s d\boldsymbol{x} - \sum_{i\geq 2}g_{s,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s d\boldsymbol{x} + \frac{1}{dt}\vartheta_s$$

Derivative i_n. The nominal interest rate derivative is very easy because it's parallel to wage inflation, except in the Phillips curve. In particular, the choice of the nominal interest rate is again a fundamentally *static* problem, even in the case with commitment. We have

$$0 = u'(c_{s,1})g_{s,1}a_1 + \mu_s g_{s,1}a_1 + \phi_{s,1}u'(c_{s,1})a_1 + \sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}g_{s,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s$$

$$+\vartheta_s \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}a_1$$

Derivative N_n. Finally, we take the derivative for aggregate labor. This is again a static problem. So, as before, we have

$$0 = \Big[u'(c_{s,1})g_{s,1}z_1A_s + \mu_s g_{s,1}z_1A_s + \phi_{s,1}u'(c_{s,1})z_1A_s + \sum_{i\geq 2}\phi_{i,s}z_iA_s\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}g_{s,i}z_iA_s\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s\Big]d\boldsymbol{x}$$

$$+\vartheta_s \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}z_1A_s d\boldsymbol{x}$$

$$-v'(N_s)\mathbf{1}'\boldsymbol{g}_s d\boldsymbol{x} - v'(N_s)\boldsymbol{\phi}_s'\mathbf{1}d\boldsymbol{x}$$

$$-\mu_sA_s + \vartheta_s\frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s(\boldsymbol{z}\cdot u'(\boldsymbol{c}_s))'\boldsymbol{g}_s d\boldsymbol{x} - v'(N_s)\right) - \frac{\epsilon}{\delta}\vartheta_s v''(N_s)N_s$$

We now summarize the resulting optimality conditions for the problem under discretion. We state these optimality conditions here for the fully discretized problem, which we have worked with thus far. For the main text, we bring these equations back to the continuous case.

We see immediately that $\vartheta_s = 0$ and

$$\boldsymbol{\phi}_s = 0$$

because the planner does not respect consumption promises from the past. These two conditions signify the lack of commitment. The optimality condition for the cross-sectional distribution still characterizes the evolution of the social lifetime value. We have

$$\begin{aligned}
0 =& u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}'_{s+1} - e^{\rho dt}\boldsymbol{\lambda}'_s}{dt} + \boldsymbol{A}\boldsymbol{\lambda}_{s+1} \\
&+ \vartheta_{s+1}N_{s+1}\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_{s+1}\boldsymbol{z}\cdot u'(\boldsymbol{c}_{s+1})
\end{aligned}$$

The optimality condition for consumption becomes

$$-\chi u''(c) = \left[u'(c) + \mu - \lambda_a + \vartheta N\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)Azu''(c)\right]g.$$

The optimality condition for monetary policy now becomes

$$0 = \left[u'(c_{s,1}) + \mu_s + \vartheta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_sN_sz_1u''(c_{s,1})\right]g_{s,1}a_1 + \sum_{i\geq 2} g_{s,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s$$

And finally, the optimality condition for aggregate economic activity becomes

$$\begin{aligned}
0 =& \left[u'(c_{s,1})g_{s,1}z_1A_s + \mu_sg_{s,1}z_1A_s + \sum_{i\geq 2} g_{s,i}z_iA_s\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s\right]d\boldsymbol{x} \\
&+ \vartheta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_sN_sz_1u''(c_{s,1})g_{s,1}z_1A_sd\boldsymbol{x} \\
&- v'(N_s) - \mu_sA_s + \vartheta_s\frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s(\boldsymbol{z}\cdot u'(\boldsymbol{c}_s))'\boldsymbol{g}_sd\boldsymbol{x} - v'(N_s)\right) - \frac{\epsilon}{\delta}\vartheta_sv''(N_s)N_s
\end{aligned}$$

Plugging in for $\vartheta_s = 0$ and $\boldsymbol{\phi}_s = 0$, we can leverage that several terms drop out and simplify to give an even sharper comparison to RANK. In RANK under discretion, the optimality condition for aggregate economic activity is given by

$$\begin{aligned}
0 =& -v'\left(\frac{Y_s}{A_s}\right)\frac{1}{A_s} + u'(Y_s) + \frac{\epsilon_s}{\delta}\vartheta_s(1+\tau^L)\frac{\epsilon_s-1}{\epsilon_s}u''(Y_s)Y_s \\
&+ \frac{\epsilon_s}{\delta}\vartheta_s\left[(1+\tau^L)\frac{\epsilon_s-1}{\epsilon_s}A_su'(Y_s) - v'\left(\frac{Y_s}{A_s}\right) - v''\left(\frac{Y_s}{A_s}\right)\frac{Y_s}{A_s}\right]\frac{1}{A_s}
\end{aligned}$$

which is unchanged from before because this is a fundamentally static optimization. *However*, as in HANK, we now have $\vartheta_s = 0$ under discretion. So we can simplify the condition as

$$Y_s = A_s^{\frac{1+\eta}{\gamma+\eta}}$$

where, importantly, the RHS is *not* equal to the natural level of output—it is if and only if there are no cost-push shocks and we have the appropriate employment subsidy. Notice that we still have the same definition of natural output, $Y_s^n = [(1+\tau^L)\frac{\epsilon_s-1}{\epsilon_s}]^{\frac{1}{\gamma+\eta}} A_s^{\frac{1+\eta}{\gamma+\eta}}$. And so we get

$$Y_s = Y_s^n \left[(1+\tau^L) \frac{\epsilon_s - 1}{\epsilon_s} \right]^{-\frac{1}{\gamma+\eta}}$$

in RANK. This tell us that commitment is only useful when there are cost-push shocks or when there is no appropriate steady state employment subsidy in place—this insight is, of course, well known from the classical analysis, which has established that the time consistency problem associated with the Phillips curve only emerges under these two conditions. See, e.g., Clarida et al. (1999) for a detailed treatment.

In HANK, on the other hand, we now have

$$0 = -\frac{1}{A} v'(N) + \int z g \lambda_a d(a,z) - \mu + \vartheta \frac{\epsilon}{\delta} \left[\frac{\epsilon - 1}{\epsilon} (1+\tau^L) A \Lambda - v'(N) - v''(N) N \right] \frac{1}{A}$$

To proceed, we first state an auxiliary result.

Lemma 36. *The social marginal value of wealth can be expressed as*

$$\lambda_a = V_a + \Big(\rho - r - \partial_t - \mathcal{A} + c_a \Big)^{-1} \Big\{ \mu c_a + \vartheta N \frac{\epsilon}{\delta} \frac{\epsilon - 1}{\epsilon} (1+\tau^L) A z u''(c) c_a \Big\}.$$

After discretizing in the space dimension, we can alternatively write this as

$$\boldsymbol{\lambda}_a = \boldsymbol{V}_a + \boldsymbol{\Gamma} \Big\{ \mu \boldsymbol{c}_a + \vartheta N \frac{\epsilon}{\delta} \frac{\epsilon - 1}{\epsilon} (1+\tau^L) A \boldsymbol{z} \cdot u''(\boldsymbol{c}) \boldsymbol{c}_a \Big\}$$

Proof. The private and social lifetime value of a household in state (a, z) are, respectively, given by

$$\rho V = (\partial_t + \mathcal{A}) V + u(c) - v(N) - \frac{\delta}{2} (\pi^w)^2$$

$$\rho \lambda = (\partial_t + \mathcal{A}) \lambda + u(c) - v(N) - \frac{\delta}{2} (\pi^w)^2 + \vartheta N \frac{\epsilon}{\delta} \frac{\epsilon - 1}{\epsilon} (1+\tau^L) A z u'(c) + \mu c$$

This should imply that we have

$$\lambda = V + \Big(\rho - \partial_t - \mathcal{A}\Big)^{-1}\Big\{\mu c + \vartheta N \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)Azu'(c)\Big\}$$

or similarly

$$\Big(\rho - \partial_t - \mathcal{A}\Big)\Big(\lambda - V\Big) = \Big\{\mu c + \vartheta N \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)Azu'(c)\Big\}$$

Differentiating with respect to a yields

$$\Big(\rho - \partial_t - \mathcal{A}\Big)\Big(\lambda_a - V_a\Big) - \Big(\frac{d}{da}\mathcal{A}\Big)\Big(\lambda - V\Big) = \mu c_a + \vartheta N \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)Azu''(c)c_a$$

or simply

$$\Big(\rho - \partial_t - \mathcal{A}\Big)\Big(\lambda_a - V_a\Big) - (r - c_a)\Big(\lambda_a - V_a\Big) = \mu c_a + \vartheta N \frac{\epsilon}{\delta}\frac{\epsilon - 1}{\epsilon}(1+\tau^L)Azu''(c)c_a.$$

Rearranging leads to the desired result. ■

Using our auxilliary Lemma 36 and plugging back into the optimality condition for aggregate economic activity in HANK, we obtain

$$\begin{aligned}0 = &-\frac{1}{A}v'(N) + \int zg\Big(V_a + \mu\kappa c_a + \kappa\vartheta N \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)Azu''(c)c_a\Big)d(a,z) - \mu \\ &+ \vartheta\frac{\epsilon}{\delta}\Big[\frac{\epsilon-1}{\epsilon}(1+\tau^L)A\Lambda - v'(N) - v''(N)N\Big]\frac{1}{A},\end{aligned}$$

where

$$\kappa = \Big(\rho - r - \partial_t - \mathcal{A} + c_a\Big)^{-1}$$

while for RANK we had

$$\begin{aligned}0 = &-\frac{1}{A}v'(N) + u'(C) + \frac{\epsilon}{\delta}\vartheta(1+\tau^L)\frac{\epsilon-1}{\epsilon}u''(C)C \\ &+ \frac{\epsilon}{\delta}\vartheta\Big[\frac{\epsilon-1}{\epsilon}(1+\tau^L)Au'(C) - v'(N) - v''(N)N\Big]\frac{1}{A}\end{aligned}$$

E.1 Proof of Proposition 2

The derivation of the targeting rule of Proposition 2 is in the main text. To obtain the CRRA representation of it, we use

$$\iint \frac{zu'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz - Y_t^{\gamma+\eta} A_t^{-(1+\eta)} = \Omega_t^{\mathrm{D}} \iint \frac{au'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz.$$

Using natural output definitions

$$\tilde{Y}_t^{\mathrm{RA}} = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) A_t^{1+\eta}\right)^{\frac{1}{\gamma+\eta}}$$

$$\tilde{Y}_t^{\mathrm{HA}} = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) A_t^{1+\eta} \iint \frac{zu'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz\right)^{\frac{1}{\gamma+\eta}}$$

we rearrange and obtain

$$\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) Y_t^{\gamma+\eta} = \frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) A_t^{1+\eta}\left(\iint \frac{zu'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz - \Omega_t^{\mathrm{D}} \iint \frac{au'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz\right)$$

or simply

$$Y_t = \tilde{Y}_t \times \underbrace{\left(\frac{\epsilon_t}{\epsilon_t - 1}\frac{1}{1+\tau^L}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Cost-Push Wedge}} \times \underbrace{\left(1 - \Omega_t^{\mathrm{D}} \frac{\iint \frac{au'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz}{\iint \frac{zu'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Redistribution}} \tag{80}$$

E.2 Proof of Proposition 3

Our expression for inflationary bias follows by plugging in the targeting rule of Proposition 2, rewritten as an expression for the aggregate labor wedge, into the Phillips curve (8). Setting $\dot{\pi}_{\mathrm{ss}}^w = 0$, this yields

$$\pi_{\mathrm{ss}}^w = -\frac{\epsilon}{\delta} A_{\mathrm{ss}} N_{\mathrm{ss}} \iint \left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L) zu'(c_{\mathrm{ss}}(a,z)) - \frac{v'(N_{\mathrm{ss}})}{A_{\mathrm{ss}}}\right) g_{\mathrm{ss}}(a,z)\, da\, dz$$

$$\pi_{\mathrm{ss}}^w = -\frac{\epsilon}{\delta} A_{\mathrm{ss}} N_{\mathrm{ss}} \iint \left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L) zu'(c_{\mathrm{ss}}(a,z)) - zu'(c_{\mathrm{ss}}(a,z)) + \Omega_{\mathrm{ss}}^{\mathrm{D}} au'(c_{\mathrm{ss}}(a,z))\right) g_{\mathrm{ss}}(a,z)\, da\, dz$$

$$\pi_{\mathrm{ss}}^w = \frac{\epsilon}{\delta} A_{\mathrm{ss}} N_{\mathrm{ss}} \left[\left(1 - \frac{\epsilon - 1}{\epsilon}(1+\tau^L)\right) \iint zu'(c_{\mathrm{ss}}(a,z)) g_{\mathrm{ss}}(a,z)\, da\, dz - \Omega_{\mathrm{ss}}^{\mathrm{D}} \iint au'(c_{\mathrm{ss}}(a,z)) g_{\mathrm{ss}}(a,z)\, da\, dz\right].$$

F Quantitative Analysis: Sensitivity, Robustness, and Further Results

F.1 Equilibrium Conditions with Shocks

We start by restating all relevant equilibrium conditions of our baseline HANK model when all three shocks are present.

The individual block is then given by

$$\begin{aligned}
\rho_t V_t(a,z) &= \partial_t V_t(a,z) + u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 + \mathcal{A}_t V_t(a,z), \\
u'(c_t(a,z)) &= \partial_a V_t(a,z) \\
\partial_t g_t(a,z) &= \mathcal{A}_t^* g_t(a,z).
\end{aligned}$$

and the aggregate block by

$$\begin{aligned}
\dot{\pi}_t^w &= \rho_t \pi_t^w + \frac{\epsilon_t}{\delta}\left[\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L) w_t \Lambda_t - v'(N_t)\right] N_t \\
Y_t &= A_t N_t \\
w_t &= A_t \\
r_t &= i_t - \pi_t \\
\pi_t &= \pi_t^w - \frac{\dot{A}_t}{A_t}.
\end{aligned}$$

Importantly, household preferences must be modified to account for time variation in the discount rate. They are now given by

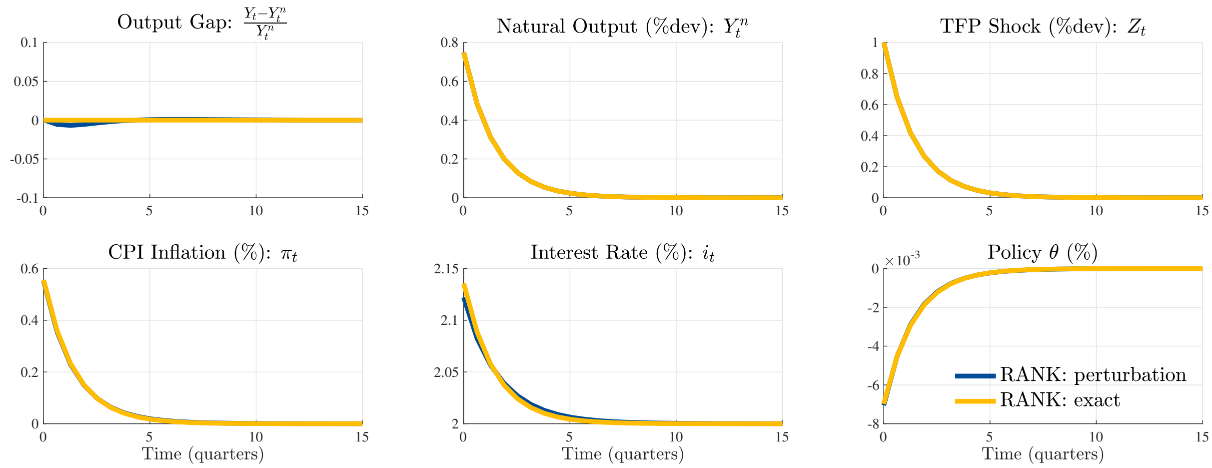
$$\int_0^\infty e^{-\int_0^t \rho_s ds}\left[u(c_t) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2\right] dt.$$

F.2 Accuracy

In this section, we report a series of numerical tests to benchmark the accuracy of our perturbation method using sequence-space Hessians. In Figure 7, we compute the transition dynamics under optimal policy in RANK in response to a TFP shock using both our perturbation method and the exact analytical solution. The Figure underscores that our first-order perturbation method is highly accurate in the case of the baseline RANK model. The remaining error in the two solutions amounts to 0.01% in the output gap or, conversely, 1bps in the optimal interest rate response.

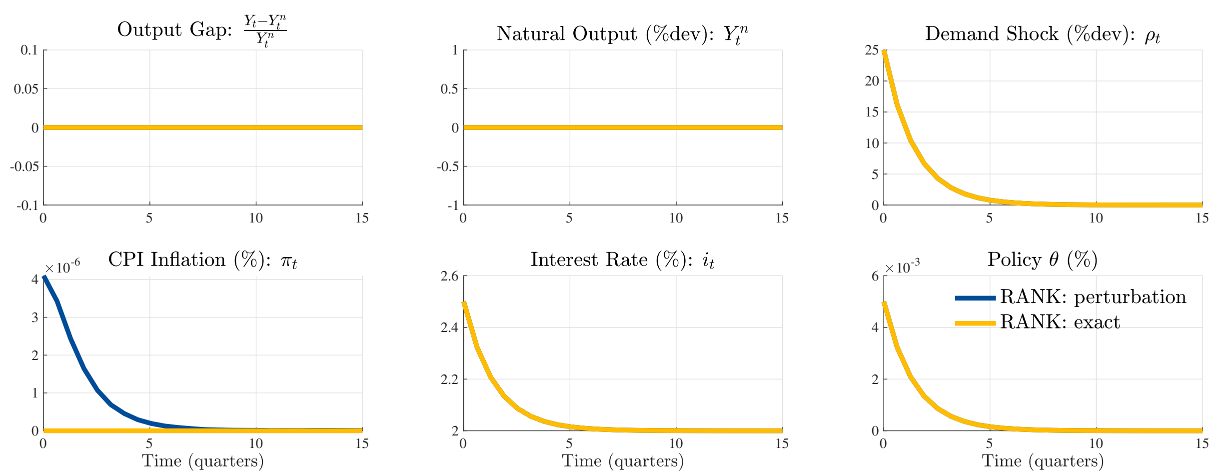
Likewise, Figure 8 reports the analogous comparison exercise for optimal policy in response to a demand shock in RANK. Here, the numerical error is even smaller. The discrepancy in optimal CPI inflation, for example, is on the order of 10^{-6}.

Figure 7. Transition Dynamics with Optimal Policy: Perturbation vs. Exact Solution



Note. Impulse responses to positive TFP under optimal monetary policy in RANK. The Figure compares the exact analytical solution of optimal policy (yellow) against our numerical perturbation approach using sequence-space Hessians (blue).

Figure 8. Transition Dynamics with Optimal Policy: Perturbation vs. Exact Solution



Note. Impulse responses to positive demand under optimal monetary policy in RANK. The Figure compares the exact analytical solution of optimal policy (yellow) against our numerical perturbation approach using sequence-space Hessians (blue).

F.3 Transition Dynamics without Optimal Policy

In this section, we present impulse response plots that display the transition dynamics of both RANK and HANK economies in response to TFP, demand, and cost-push shocks without optimal policy interventions. We model monetary policy instead as following a Taylor rule, with

$$i_t = r^{ss} + \lambda_\pi \pi_t, \tag{81}$$

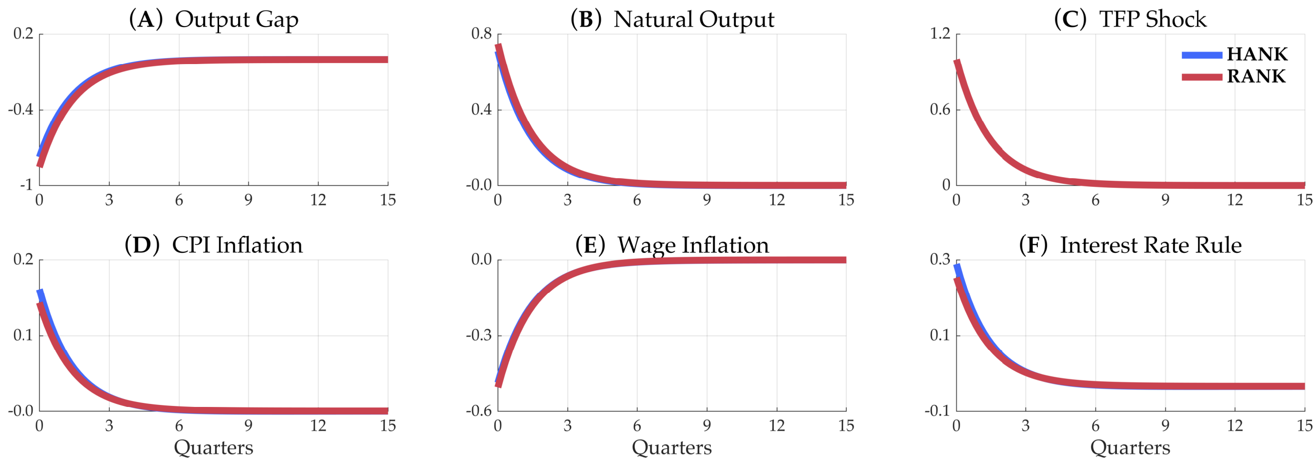


Figure 9. Transition Dynamics under Taylor Rule: TFP Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (81) and is not set optimally. The cost-push shock is modeled as an increase in labor union's desired wage mark-up. The shock is initialized at $\epsilon_0 = 11$ and mean-reverts to its steady state value $\epsilon = 10$, with a half-life of 2 quarters.

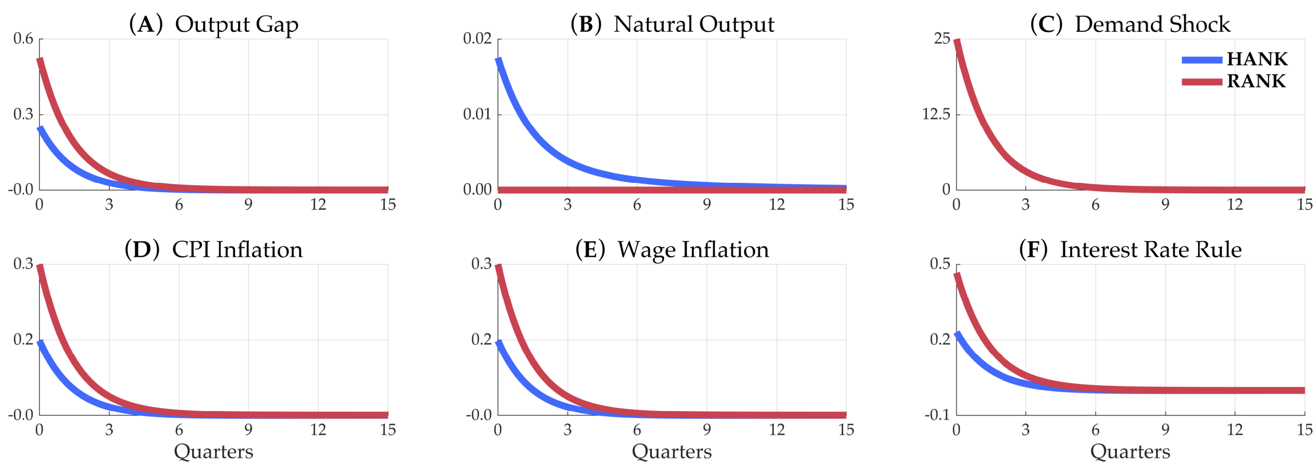


Figure 10. Transition Dynamics under Taylor Rule: Demand Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (81) and is not set optimally. The cost-push shock is modeled as an increase in labor union's desired wage mark-up. The shock is initialized at $\epsilon_0 = 11$ and mean-reverts to its steady state value $\epsilon = 10$, with a half-life of 2 quarters.

where we calibrate $\lambda_\pi = 1.5$.

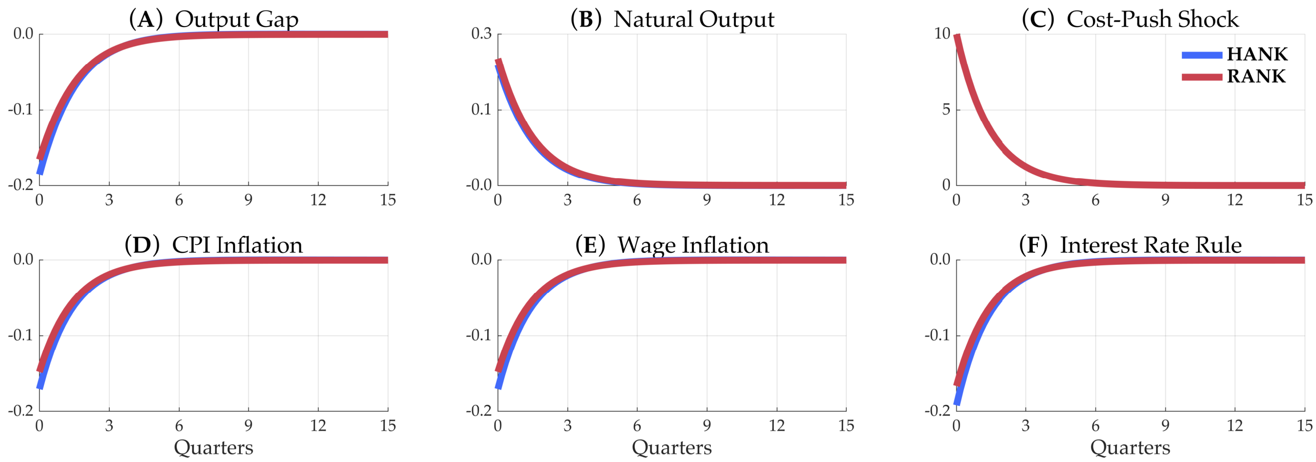


Figure 11. Transition Dynamics under Taylor Rule: Cost-Push Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (81) and is not set optimally. The cost-push shock is modeled as an increase in labor union's desired wage mark-up. The shock is initialized at $\epsilon_0 = 11$ and mean-reverts to its steady state value $\epsilon = 10$, with a half-life of 2 quarters.

F.4 Sensitivity Analysis: Low Inequality

In this section, we explore the implications of inequality for optimal stabilization policy. In our benchmark calibration in Section 5, we set $z^L = 0.8$ and $z^H = 1.2$. The RANK benchmark can be viewed instead as the limit where $z^L, z^H \to \bar{z} = 1$.

We again compute the impulse responses to TFP, demand, and cost-push shocks under optimal policy. However, we solve an alternative calibration with less uninsurable earnings risk, setting $z^L = 0.95$ and $z^H = 1.05$. This calibration is between the RANK and HANK benchmarks of Section 5. We use this robustness exercise to confirm that the hump-shapes of the IRFs under optimal policy in Section 5 are driven by distributional considerations. We indeed find that less uninsurable risk and, consequently, lower cross-sectional consumption dispersion lead to less pronounced hump-shapes in the optimal policy IRFs.

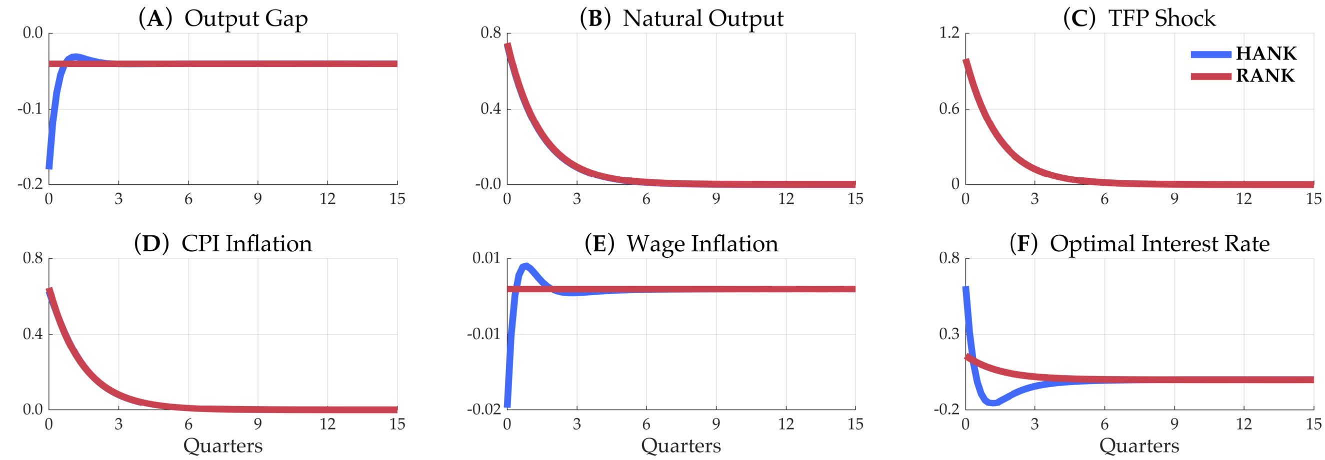


Figure 12. Optimal Policy Transition Dynamics with Low Earnings Rrisk: TFP Shock

Note. Impulse responses to positive TFP shock in both RANK (yellow) and HANK (blue) models under optimal monetary stabilization policy, using an alternative calibration with less uninsurable earnings risk with $z^{\mathrm{L}} = 0.95$ and $z^{\mathrm{H}} = 1.05$. The initial shock is 1% of steady state TFP and mean-reverts with a half-life of 2 quarters.

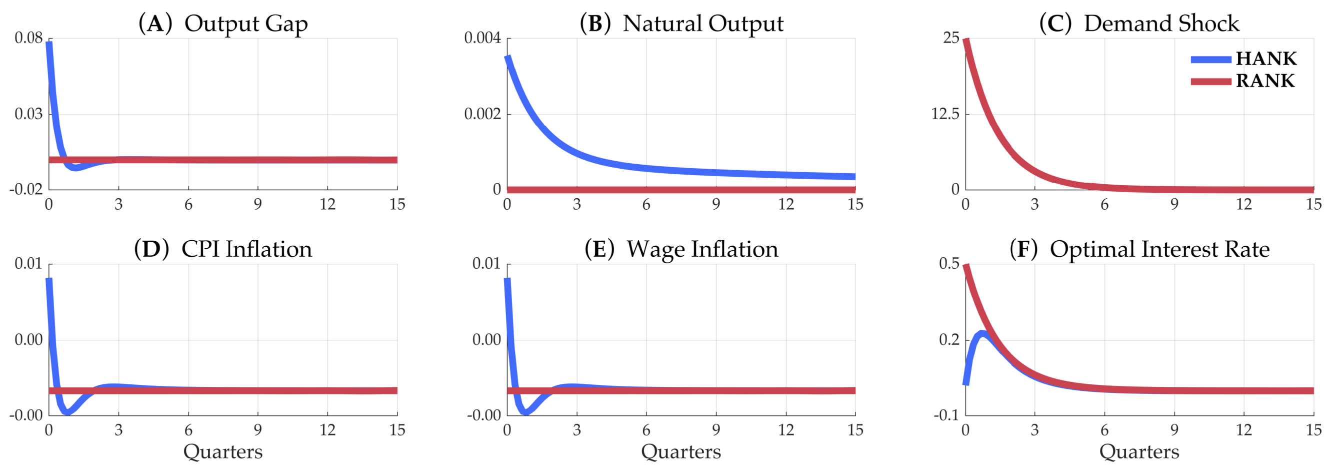


Figure 13. Optimal Policy Transition Dynamics with Low Earnings Rrisk: Demand Shock

Note. Impulse responses to positive discount rate shock in both RANK (yellow) and HANK (blue) models under optimal monetary stabilization policy, using an alternative calibration with less uninsurable earnings risk with $z^{\mathrm{L}} = 0.95$ and $z^{\mathrm{H}} = 1.05$. 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 2 quarters.

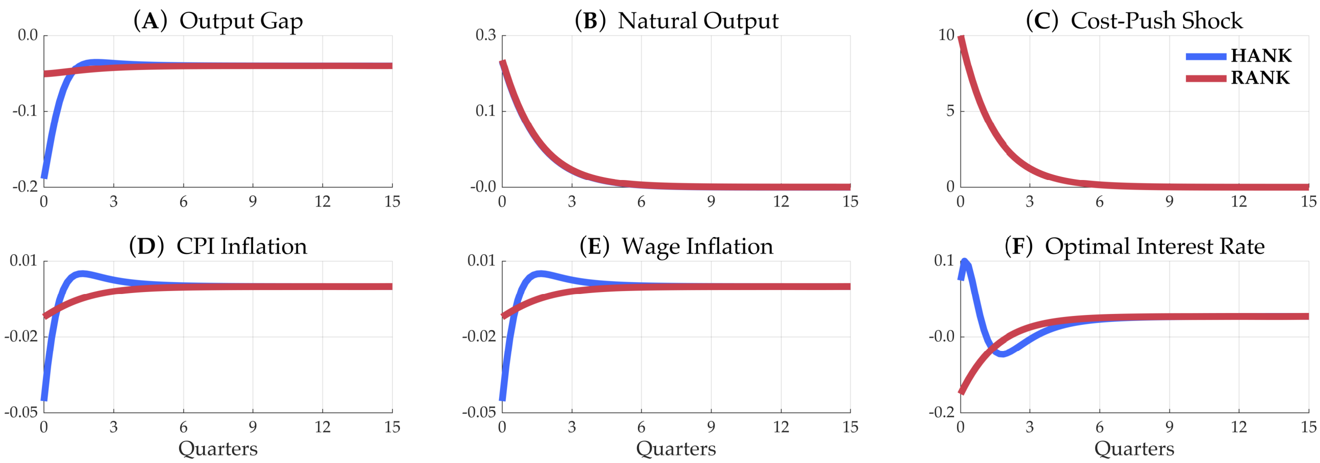


Figure 14. Optimal Policy Transition Dynamics with Low Earnings Rrisk: Cost-Push Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models under optimal monetary stabilization policy, using an alternative calibration with less uninsurable earnings risk with $z^{L} = 0.95$ and $z^{H} = 1.05$. The cost-push shock is modeled as an increase in labor union's desired wage mark-up. The shock is initialized at $\epsilon_0 = 11$ and mean-reverts to its steady state value $\epsilon = 10$, with a half-life of 2 quarters.