Online Appendix (not for publication)

A Additional Model Details and Derivations

A.1 Labor Unions and Wage Rigidity

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

$$\dot{a}_t = r_t a_t + z_t \frac{1}{P_t} \int_0^1 W_{k,t} n_{k,t} dk + \tau_t(z_t) - c_t. \tag{54}$$

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

$$N_{k,t} = \left(\frac{W_{k,t}}{W_t}\right)^{-\epsilon_t} N_t \tag{55}$$

$$W_t = \left(\int_0^1 W_{k,t}^{1-\epsilon_t} dk \right)^{\frac{1}{1-\epsilon_t}} \tag{56}$$

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

$$\pi^w_{k,t} = \frac{\dot{W}_{k,t}}{W_{k,t}}. \tag{57}$$

The union's adjustment cost is directly passed to union members as a quadratic utility cost, so households' instantaneous flow utility is formally given by

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

where $v(\cdot)$ captures pure disutility from working and δ modulates the strength of the wage rigidity.[48] The representation in the main text, i.e., equation (4), is valid in any equilibrium that features symmetric unions, which we assume.

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. That is, union k solves

$$\max_{\pi^w_{k,t}} \int_0^\infty e^{-\int_0^t \rho_s ds}\left(\iint \left[u(c_t(a,z;W_{k,t})) - v\left(\int_0^1 n_{k,t}\, dk\right) - \frac{\delta}{2}\int_0^1 (\pi^w_{k,t})^2\, dk\right] g_t(a,z)\, da\, dz\right) dt, \tag{58}$$

subject to equations (55) and (57). 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 (58). However, since union k is small, it takes as given all macroeconomic aggregates, including the cross-sectional household distribution.

Solving the union problem. To solve the union's problem we associate it with the Lagrangian

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

[48] 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.

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

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

$$0 = -\delta \pi^w_{k,t} + \mu_t W_{k,t},$$

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

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

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. In particular, when equilibrium is initialized at a symmetric nominal wage distribution $\{W_{k,0}\}$, then the wage policies that result from the union's problem maintain symmetry of equilibrium. That is, wages and labor allocations remain equalized across unions, with $W_{k,t} = W_t$ and $N_{k,t} = N_t$. In such a symmetric equilibrium, the non-linear New Keynesian wage Phillips curve is then as in the main text.

A.2 Fiscal Rebates

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. 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

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

A.3 Natural Output and the Flexible Wage Limit

We define natural output as the output that obtains in the limit of flexible wages, i.e., as $\delta \to 0$. With isoelastic (CRRA) preferences $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$ and $v(n) = \frac{1}{1+\eta} n^{1+\eta}$, natural output in HANK is given by

$$\tilde{Y}_t = \left(\frac{\epsilon_t - 1}{\epsilon_t} (1 + \tau^L) A_t^{1+\eta} \iint \frac{z u'(c_t(a,z))}{u'(C_t)} g_t(a,z)\, da\, dz \right)^{\frac{1}{\gamma+\eta}}, \tag{59}$$

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

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 flexible wage allocation, and we obtain natural output from this equation.

A.4 Competitive Equilibrium and Implementability

To conclude our discussion of the model details, we now state formally the implementability conditions for the Ramsey problem in continuous time. As part of our discussion, we also provide additional details on the generator $\mathcal{A}_t$ and its adjoint $\mathcal{A}_t^*$, which we use in the main text. Finally, in Section A.5, we develop a discretized representation of these implementability conditions, which we leverage in our proofs below.

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:

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

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

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

where $\mathcal{A}_t$ 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

$$\mathcal{A}_t f_t(a,z) = \Big(r_t a + z w_t N_t - c_t(a,z) \Big) \partial_a f_t(a,z) + \mathcal{A}_z f_t(a,z), \tag{60}$$

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

$$\mathcal{A}_t^* f_t(a,z) = -\partial_a\Big[\Big(r_t a + z w_t N_t - c_t(a,z)\Big) f_t(a,z)\Big] + \mathcal{A}_z^* f_t(a,z), \tag{61}$$

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:

$$\begin{aligned}
\dot{\pi}_t^w &= \rho_t \pi_t^w + \frac{\epsilon_t}{\delta} \iint n_t \left(\frac{\epsilon_t - 1}{\epsilon_t} (1+\tau^L) w_t z u'(c_t) - v'(n_t) \right) g_t(a,z)\, da\, dz \\
Y_t &= A_t N_t \\
w_t &= A_t \\
r_t &= i_t - \pi_t \\
\pi_t &= \pi_t^w - \frac{\dot{A}_t}{A_t}.
\end{aligned}$$

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

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

We conclude this subsection by characterizing the operator $\mathcal{M}_t(a,z)$, which is an important input in the targeting rules we present in the main text (and in the proofs below). In particular, the operator admits the representation

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

where 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).

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.

A.5 Discretized Competitive Equilibrium Conditions

We now develop a discretized representation of competitive equilibrium and the associated implementability conditions for the Ramsey problem. This discretized representation will elucidate how boundary conditions are treated formally by the Ramsey planner. In particular, we leverage this representation to explicitly account for households' borrowing constraint when deriving our proofs below.

For any function $c_t(a,z)$, we discretize both in the individual state space (a,z) and in time t. We denote this discretization by $\boldsymbol{c}_n$ 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 and date t_n. 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 follow Achdou et al. (2022) 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. The following Lemma summarizes the discretized competitive equilibrium conditions of our model, using a finite-difference discretization given a policy path $\boldsymbol{i} = \{i_n\}_{n \geq 0}$. The proof follows along the lines of Achdou et al. (2022) and Schaab and Zhang (2022), 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 13. *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

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

Finally, for the resource constraint we simply have

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

and for the Phillips curve

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

and we have already used $c_{n,1} = i_n a_1 - \pi^w_n a_1 + \frac{A_{n+1} - A_n}{dt A_n} a_1 + z_1 A_n N_n$.

In this Lemma, we denote by $\boldsymbol{D}_a$ the finite-difference matrix that discretizes the partial derivative operator ∂_a. We also denote by $\boldsymbol{A}^z$ the (finite-difference) matrix that discretizes the operator $\mathcal{A}^z$ associated with the earnings process. Finally, $d\boldsymbol{x}$ denotes the integration measure of households. See Schaab and Zhang (2022) for details.

Crucially, the discretized system of equations in the above Lemma properly accounts for the household borrowing constraint, leveraging results from Achdou et al. (2022). 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.

B Appendix for Section 4

We invert our presentation of proofs and formal derivations for Sections 3 and 4. In this Appendix, we start by setting up and characterizing the standard Ramsey problem, which is an instructive building block for the proofs that follow.

In particular, we state the continuous-time Ramsey problem in Section B.1 and present an illustrative but heuristic derivation of its first-order conditions. To formally account for boundary conditions, we then introduce and characterize the discretized standard Ramsey problem in Section B.2, leveraging the discretized representation of equilibrium from Appendix A.5.[49]

B.1 Standard Ramsey Problem in Continuous Time

In this section, we restate for convenience the standard Ramsey problem in continuous time and develop a heuristic derivation of its optimality conditions. We defer a formal treatment of boundary conditions to Appendix B.2. To adopt more compact notation, 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 $\partial_t c_t(a,z)$ becomes c_t.

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

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

[49] 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 of our paper. 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.

Heuristic derivation. We provide an illustrative but heuristic derivation of the optimality conditions, abstracting from formally taking into account boundary conditions. This derivation is valuable because it is relatively brief and accessible. The proofs that follow become more complex only insofar as they formally take into account boundary conditions.

The following auxilliary results will be helpful. Integrating various partial derivatives in the above Lagrangian by parts, we have

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

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

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

Next, for the adjoint, we have

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

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

$$\int_0^\infty e^{-\rho t} \int \phi \mathcal{A} V da\, dz\, dt = \int_0^\infty e^{-\rho t} \iint V \mathcal{A}^* \phi da\, dz\, dt.$$

Finally, for the consumption FOC, we simply have

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

where we also drop boundary terms.

The functional Lagrangian can thus be rewritten as

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

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

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

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:

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

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,

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

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, i.e., abstracting from boundary conditions.

B.2 Discretized Standard Ramsey Problem

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. To that end, we work with the discretized representation of equilibrium developed in Appendix A.5.

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

$$
\begin{aligned}
L^{\mathrm{SP}}(\boldsymbol{g}_0) = & \min_{\{\boldsymbol{\phi}_n, \boldsymbol{\chi}_n, \boldsymbol{\lambda}_n, \mu_n, \theta_n\}} \max_{\{\boldsymbol{V}_n, \boldsymbol{c}_{n,[2:J]}, \boldsymbol{g}_n, \pi_n^w, N_n, i_n\}} \sum_{n=0}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ \\
& + 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}' \boldsymbol{g}_t - v(N_n) \mathbf{1}' \boldsymbol{g}_t - \frac{\delta}{2} (\pi_n^w)^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_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 \right] \\
& + \boldsymbol{\phi}_n' \boldsymbol{A}^z \boldsymbol{V}_n + \sum_{i \geq 2} \phi_{i,n} \left(i_n a_i - \pi_n^w a_i + \frac{A_{n+1} - A_n}{dt A_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_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 \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] \\
& + \theta_n \left[-\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 \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.5, we fix from the beginning that unemployed households at the borrowing constraint always consume their income, that is

$$c_{n,1} = i_n a_1 - \pi_n^w a_1 + \frac{A_{n+1} - A_n}{dt A_n} a_1 + z_1 A_n N_n$$

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 $\dot{\pi}_t^w$. These assumptions also correspond to the appropriate stencils we use numerically to implement our results.

We also want to echo Achdou et al. (2022) 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, $-\boldsymbol{D}_a(\boldsymbol{s} \cdot \boldsymbol{g})$.

B.3 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 14. *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

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

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

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

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

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

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

Lemma 16. *The correct adjoint operation, i.e., the one we use to define* $\mathcal{A}^* \approx \boldsymbol{A}'$*, is given by*

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

In particular, we have

$$\begin{aligned}\boldsymbol{\lambda}'(\boldsymbol{A}^a)'\boldsymbol{g} &= \boldsymbol{\lambda}'\boldsymbol{D}_a'(\boldsymbol{s}\cdot\boldsymbol{g})\\ &= (\boldsymbol{s}\cdot\boldsymbol{g})'\boldsymbol{D}_a\boldsymbol{\lambda}\\ &= \boldsymbol{g}'(\boldsymbol{s}\cdot\boldsymbol{D}_a)\boldsymbol{\lambda}\\ &= (\boldsymbol{D}_a\boldsymbol{\lambda})'(\boldsymbol{s}\cdot\boldsymbol{g}).\end{aligned}$$

Lemma 17. *We can "integrate by parts" the FOC term in the Lagrangian to arrive at*

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

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 18. *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

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

as well as

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

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

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

Proof. We have

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

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

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

Lastly,

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

And so we get

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

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

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

B.4 Proof of Proposition 4

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^{\mathrm{SP}}(\boldsymbol{g}_0) = \min_{\{\boldsymbol{\phi}_n,\boldsymbol{\chi}_n,\boldsymbol{\lambda}_n,\mu_n,\theta_n\}} \max_{\{\boldsymbol{V}_n,\boldsymbol{c}_{n,[2:J]},\boldsymbol{g}_n,\pi_n^w,N_n,i_n\}} \sum_{n=0}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ & 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}' \boldsymbol{g}_n \\
& + \mu_n \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}' \boldsymbol{g}_n - v(N_n)\mathbf{1}'\boldsymbol{g}_n - \frac{\delta}{2}(\pi_n^w)^2\mathbf{1}'\boldsymbol{g}_n \\
& - \frac{\boldsymbol{\phi}_n' - e^{\rho dt}\boldsymbol{\phi}_{n-1}'}{dt}\boldsymbol{V}_n + \boldsymbol{\phi}_n'\left[-\rho \boldsymbol{V}_n + 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 \right] \\
& + \boldsymbol{\phi}_n' \boldsymbol{A}^z \boldsymbol{V}_n + \frac{1}{da}\boldsymbol{V}_n'(\boldsymbol{\phi}_n \cdot \boldsymbol{D}_a)' \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} \\
& + \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{\lambda}_n' - e^{\rho dt}\boldsymbol{\lambda}_{n-1}'}{dt}\boldsymbol{g}_n + \boldsymbol{\lambda}_n'(\boldsymbol{A}^z)'\boldsymbol{g}_n \\
& + \frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\lambda}_t)'\left[\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 \boldsymbol{g}_n \right] \Bigg\} d\boldsymbol{x} \\
& - \mu_n A_n N_n \\
& + \frac{\theta_n - e^{\rho dt}\theta_{n-1}}{dt}\pi_n^w + \theta_n \rho \pi_n^w - \theta_n \frac{\epsilon}{\delta} v'(N_n) N_n \\
& + \theta_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_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} d\boldsymbol{x} \Bigg\} dt \\
& - e^{\rho dt}\boldsymbol{\phi}_{-1}'\boldsymbol{V}_0 d\boldsymbol{x} + e^{\rho dt}e^{-\rho t_N}\boldsymbol{\phi}_{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{\lambda}_{-1}'\boldsymbol{g}_0 d\boldsymbol{x} - e^{\rho dt}e^{-\rho t_N}\boldsymbol{\lambda}_{N-1}'\boldsymbol{g}_N d\boldsymbol{x} \\
& + e^{\rho dt}\theta_{-1}\pi_0^w - e^{\rho dt}e^{-\rho t_N}\theta_{N-1}\pi_N^w.
\end{aligned}
$$

In the spirit of Marcet and Marimon (2019), we have reordered the forward-looking constraints—this corresponds to summation (integration) by parts. The resulting "boundary" terms in the last few lines of the above Lagrangian are the key objects at the heart of the time-0 problem, which we discuss in Section 4.

To conclude our proof of Proposition 4, we now take derivatives and characterize the necessary first-order conditions. In particular, we do so for $n \geq 1$ precisely in order to avoid the boundary terms that give rise to the time-0 problem. In the continuous time limit as $dt \to 0$, these terms give rise to the initial conditions $\phi_0(a,z) = 0$ and $\theta_0 = 0$, as we explain in our statement of Proposition 4 and the discussion in the main text. This follows straightforwardly from basic calculus of variations (see e.g. Kamien and Schwartz, 2012). We revisit these boundary terms below when we prove the timeless property of the timeless Ramsey plans.

Derivative V_n. We have

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

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

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

Derivative $\boldsymbol{g}_n$. We have

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

Now we work out the remaining derivative,

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

Thus, we have

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

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

$$\begin{aligned} 0 =& u'(c_{n,i})g_{n,i} + \mu_n g_{n,i} + u'(c_{n,i})\phi_{n,i} + u''(c_{n,i})\chi_{n,i} + \theta_n N_n \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_n z_i u''(c_{n,i})g_{n,i} \\ &+ \frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n'(\boldsymbol{\phi}_n \cdot \boldsymbol{D}_a)\boldsymbol{V}_n\right] + \frac{d}{dc_{n,i}}\left[\frac{1}{da}\boldsymbol{s}_n' \cdot (\boldsymbol{D}_a\boldsymbol{\lambda}_n)'\boldsymbol{g}_n\right]. \end{aligned}$$

Working out the remaining derivatives, we have

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

And similarly,

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

Thus, we have

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

Derivative π_n^w. We have

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

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

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

$$+ \frac{\theta_n - e^{\rho dt} \theta_{n-1}}{dt} + \rho \theta_n.$$

Alternatively, we have

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

Thus, we have

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

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

$$- \sum_{i \geq 2} \phi_{i,n} a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n d\boldsymbol{x} - \sum_{i \geq 2} g_{n,i} a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{\lambda}_n d\boldsymbol{x} + \frac{\theta_n - e^{\rho dt} \theta_{n-1}}{dt} + \rho \theta_n.$$

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

$$
\begin{aligned}
0 =& u'(c_{n,1}) g_{n,1} a_1 + \mu_n g_{n,1} a_1 + \phi_{n,1} u'(c_{n,1}) a_1 + \sum_{i\geq 2} \phi_{i,n} a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n + \sum_{i\geq 2} g_{n,i} a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{\lambda}_n \\
&+ \theta_n \frac{\epsilon}{\delta} \frac{\epsilon - 1}{\epsilon} (1+\tau^L) A_n N_n z_1 u''(c_{n,1}) g_{n,1} a_1.
\end{aligned}
$$

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

$$
\begin{aligned}
0 =& \left[u'(c_{n,1}) g_{n,1} z_1 A_n + \mu_n g_{n,1} z_1 A_n + \phi_{n,1} u'(c_{n,1}) z_1 A_n + \sum_{i\geq 2} \phi_{i,n} z_i A_n \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{V}_n + \sum_{i\geq 2} g_{n,i} z_i A_n \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{\lambda}_n \right] d\boldsymbol{x} \\
&+ \theta_n \frac{\epsilon}{\delta} \frac{\epsilon - 1}{\epsilon} (1+\tau^L) A_n N_n z_1 u''(c_{n,1}) g_{n,1} z_1 A_n d\boldsymbol{x} \\
&- v'(N_n) \boldsymbol{1}' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \boldsymbol{\phi}_n' \boldsymbol{1} d\boldsymbol{x} \\
&- \mu_n A_n + \theta_n \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1+\tau^L) A_n (\boldsymbol{z} \cdot u'(\boldsymbol{c}_n))' \boldsymbol{g}_n d\boldsymbol{x} - v'(N_n) \right) - \frac{\epsilon}{\delta} \theta_n v''(N_n) N_n.
\end{aligned}
$$

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.

B.5 Stationary Ramsey Plan and Proof of Proposition 5

We now formally state the discretized characterization 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 20. (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*

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

and for the distribution

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

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

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

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

$$0 = \left(u'(c_1) + \mu - \frac{1}{da}\boldsymbol{D}_{a,[1,:]}\boldsymbol{\lambda} + \theta\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)Nz_1 u''(c_1)\right) g_1 a_1 + \sum_{i\geq 1}\phi_i a_i u'(c_i) + \sum_{i \geq 1} g_i a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}.$$

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

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

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

$$0 = \left[\left(u'(c_1) + \mu - \frac{1}{da}\boldsymbol{D}_{a,[1,:]}\boldsymbol{\lambda} + \theta\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)Nz_1 u''(c_1)\right) g_1 z_1 + \sum_{i\geq 1}\phi_i z_i u'(c_i) + \sum_{i \geq 1} g_i z_i \frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}\right] d\boldsymbol{x}$$
$$- v'(N) - v'(N)\boldsymbol{\phi}'\boldsymbol{1}d\boldsymbol{x} - \mu + \theta\frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L)(\boldsymbol{z}\cdot u'(\boldsymbol{c}))'\boldsymbol{g}d\boldsymbol{x} - v'(N)\right) - \frac{\epsilon}{\delta}\theta v''(N)N.$$

This representation follows from setting all equilibrium objects to constants, e.g., $\theta_n = \theta$. This Lemma states necessary conditions that any stationary Ramsey plan must satisfy. It does not speak to convergence to a stationary Ramsey plan. 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.

From the stationary counterpart of equation (30), 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 (30). It implies that the "births" and "deaths" of distributional penalties must average out to zero in a stationary Ramsey plan. In additional derivations available upon request, 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.

Proof of Proposition 5. Note that in any stationary equilibrium we must have $\boldsymbol{\phi}'\mathbf{1}d\boldsymbol{x} = 1$, which we show below. Now notice that

$$\frac{1 - e^{\rho dt}}{dt}\theta + \rho\theta \to \frac{1 - 1 - \rho dt}{dt}\theta + \rho\theta = 0$$

in the limit as $dt \to 0$. Therefore, we must have $\pi^w = \pi = 0$ at the stationary Ramsey plan.

B.6 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.[50] 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*

$$\max_{\{i_t\}} L^{\mathrm{TD}}(g_{ss}, \phi_{ss}, \theta_{ss}),$$

where $L^{TD}(g_{ss}, \phi_{ss}, \theta_{ss})$ denotes the timeless dual Lagrangian, given an initial distribution g_{ss} as well as initial promises ϕ_{ss} and θ_{ss}. The Lagrangian is defined as

$$L^{\mathrm{TD}}(g_{ss}, \phi_{ss}, \theta_{ss}) = \int_0^\infty e^{-\rho t} \iint \left[u(c_t(a,z)) - v(N_t) - \frac{\delta}{2}(\pi_t^w)^2 \right] g_t(a,z)\, da\, dz\, dt + \mathcal{T}(\phi_{ss}, \theta_{ss}), \quad (62)$$

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

Proposition 21. (**Timeless Ramsey Problem Resolves Time-0 Problem under Primal and Dual**) *Optimal policy under the timeless primal and dual Lagrangians resolves the time-0 problem. That is,*

$$\frac{d}{d\boldsymbol{i}} L^{\mathrm{TD}}\Big(g_{ss}, \phi_{ss}, \theta_{ss}, \boldsymbol{i}_{ss}, \boldsymbol{Z}_{ss}\Big) = \frac{d}{d\boldsymbol{i}} L^{\mathrm{TP}}\Big(g_{ss}, \phi_{ss}, \theta_{ss}, \boldsymbol{i}_{ss}, \boldsymbol{Z}_{ss}\Big) = 0. \quad (63)$$

[50] 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.

Proposition 21 is a slightly more general variant of Proposition 6 in the main text. In the main text, we exclusively discuss the primal form of the Ramsey problem. In this Appendix, we leverage the duality between primal and dual because some results and proofs are easier to derive in one or the other. Proposition 21 establishes that both the primal and dual timeless Ramsey problems resolve the time-0 problem. We now prove this result.

B.7 Proof of Proposition 6

Our goal is to show that

$$\frac{dL^{\mathrm{TP}}(\boldsymbol{g}_{\mathrm{ss}}, \boldsymbol{\phi}_{\mathrm{ss}}, \theta_{\mathrm{ss}}, \boldsymbol{i}_{\mathrm{ss}}, \boldsymbol{Z}_{\mathrm{ss}})}{d\boldsymbol{i}} = F\Big(\boldsymbol{g}_{\mathrm{ss}}, \boldsymbol{\phi}_{\mathrm{ss}}, \theta_{\mathrm{ss}}, \boldsymbol{i}_{\mathrm{ss}}, \boldsymbol{Z}_{\mathrm{ss}}\Big) = 0. \tag{64}$$

We proceed as follows: First, instead of working with the timeless primal Lagrangian, we leverage the observation that $\frac{d}{d\boldsymbol{i}}L^{\mathrm{TP}} = 0$ if and only if the analogous perturbation for the timeless *dual* Lagrangian is 0, i.e., $\frac{d}{d\boldsymbol{i}}L^{\mathrm{TD}} = 0$. We make this point in our discussion of the dual approach below, noting the generic duality between primal and dual representations of the Ramsey problem. Second, we will prove that this perturbation is 0 for a given $\frac{d}{di_k}$, i.e., for a one-time perturbation in the interest rate at time k. We can then simply "stack" up to arrive at any perturbation $\frac{d}{d\boldsymbol{i}}$.

For our baseline HANK model, $\frac{dL^{\mathrm{TD}}}{di_k}$ takes the form

$$0 = \frac{d}{di_k}\Bigg\{\sum_{n=0}^{\infty} e^{-\rho t}\Bigg\{u\begin{pmatrix} i_n a_1 - \pi_n^w a_1 + \frac{A_{n+1}-A_n}{dtA_n}a_1 + z_1 A_n N_n \\ \boldsymbol{c}_{n,[2:J]}\end{pmatrix}' \boldsymbol{g}_n - v(N_n)\mathbf{1}'\boldsymbol{g}_n - \frac{\delta}{2}(\pi_n^w)^2\mathbf{1}'\boldsymbol{g}_n\Bigg\}d\boldsymbol{x}dt$$
$$\underbrace{+\frac{1}{dt}e^{\rho dt}\boldsymbol{\phi}'\boldsymbol{V}_0 d\boldsymbol{x} - \frac{1}{dt}e^{\rho dt}\theta\pi_0^w}_{\text{Timeless Penalties}}\Bigg\}\Bigg|_{\boldsymbol{g}_{\mathrm{ss}},\boldsymbol{\phi}_{\mathrm{ss}},\theta_{\mathrm{ss}},\boldsymbol{i}_{\mathrm{ss}},\boldsymbol{Z}_{\mathrm{ss}}}$$

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

$$0 = \sum_{n=0}^{\infty} e^{-\rho t}\Bigg\{\Big[u(\boldsymbol{c}_n) - v(N_n) - \frac{\delta}{2}(\pi_n^w)^2\Big]'\frac{d\boldsymbol{g}_t}{di_k} + (\boldsymbol{g}_n \cdot u'(\boldsymbol{c}_n))'\frac{d\boldsymbol{c}_n}{di_k} - (v'(N_n)\mathbf{1} + \delta\pi_n^w\mathbf{1})'\boldsymbol{g}_n\frac{dN_n}{di_k}\Bigg\}dt$$
$$+\frac{1}{dt}e^{\rho dt}\boldsymbol{\phi}'\frac{d\boldsymbol{V}_0}{di_k} - \frac{1}{dt}e^{\rho dt}\theta\frac{d\pi_0^w}{di_k}\frac{1}{d\boldsymbol{x}},$$

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

Our proof strategy will be to add 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

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

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

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

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 i_k, we obtain

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

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

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

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

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

Differentiating, we obtain

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

Evaluate at stationary Ramsey plan. Crucially, each of our five auxilliary equations must necessarily also hold when evaluated around the 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 $\frac{dL^{\text{TD}}}{di_k}$ which we started out with, where we now also evaluate the latter at the stationary Ramsey plan. This yields

$$
\begin{aligned}
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}{di_k} + (\boldsymbol{g} \cdot u'(\boldsymbol{c}))' \frac{d\boldsymbol{c}_n}{di_k} - (v'(N)\mathbf{1} + \delta \pi^w \mathbf{1})' \boldsymbol{g} \frac{dN_n}{di_k} \\
& + (\boldsymbol{\phi} \cdot u'(\boldsymbol{c})) \frac{d\boldsymbol{c}_n}{di_k} - \boldsymbol{\phi}' \mathbf{1} \left(v'(N) \frac{dN_n}{di_k} + \delta \pi_n^w \frac{d\pi_n^w}{di_k} \right) \\
& + \boldsymbol{\phi}' \boldsymbol{A}^z \frac{d\boldsymbol{V}_n}{di_k} + \frac{1}{da} (\boldsymbol{\phi} \cdot \boldsymbol{D}_a \boldsymbol{V})' \frac{d\boldsymbol{s}_n}{di_k} + \frac{1}{da} \boldsymbol{\phi}' \boldsymbol{s} \cdot \boldsymbol{D}_a \frac{d\boldsymbol{V}_n}{di_k} \\
& + \left(\boldsymbol{\chi}_{[2:J]} \cdot u''(\boldsymbol{c}_{[2:J]}) \right)' \frac{d\boldsymbol{c}_{n,[2:J]}}{di_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}{di_k} \\
& - \rho \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_n}{di_k} + \boldsymbol{\lambda}' (\boldsymbol{A}^z)' \frac{d\boldsymbol{g}_n}{di_k} + (\boldsymbol{g} \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{s}_n}{di_k} + (\boldsymbol{s} \cdot \boldsymbol{D}_a \boldsymbol{\lambda})' \frac{d\boldsymbol{g}_n}{di_k} \\
& + \mu \frac{1}{d\boldsymbol{x}} A \frac{dN_n}{di_k} - \mu \boldsymbol{c}' \frac{d\boldsymbol{g}_n}{di_k} - \mu \boldsymbol{g}' \frac{d\boldsymbol{c}_n}{di_k} \\
& + \theta \frac{\epsilon}{\delta} \left(\frac{\epsilon - 1}{\epsilon} (1 + \tau^L) A \left((\boldsymbol{z} \cdot u'(\boldsymbol{c}))' \frac{d\boldsymbol{g}_n}{di_k} + (\boldsymbol{z} \cdot u''(\boldsymbol{c}) \cdot \boldsymbol{g})' \frac{d\boldsymbol{c}_n}{di_k} \right) d\boldsymbol{x} - v''(N) \frac{dN_n}{di_k} \right) N_n \\
& + \theta \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}{di_k} \Bigg\} dt \\
& + e^{\rho dt} \boldsymbol{\phi}' \frac{d\boldsymbol{V}_0}{di_k} - e^{\rho dt} \theta \frac{d\pi_0^w}{di_k} \frac{1}{d\boldsymbol{x}} \\
& - e^{\rho dt} \boldsymbol{\phi}' \frac{d\boldsymbol{V}_0}{di_k} + e^{\rho dt} \boldsymbol{\lambda}' \frac{d\boldsymbol{g}_0}{di_k} + e^{\rho dt} \theta \frac{d\pi_0^w}{di_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}{di_k},
\end{aligned}
$$

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 $\frac{d\boldsymbol{g}_0}{di_k} = 0$ 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

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

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

$$\frac{dL^{\text{TD}}}{di_k} = 0$$

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}—and consequently also the timeless primal Lagrangian L^{TP}—indeed resolves the time-0 problem. This proof demonstrates that our timeless Ramsey approach formalizes Woodford (1999)'s timeless perspective in our HANK economy.

B.8 Proof of Proposition 7

From equation (33), we have that

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

Combining equation (32) and the definition of ξ_t^{HTM}, we have

$$\begin{aligned} 0 = &\iint zu'(c_t(a,z))\, g_t(a,z)\, dadz - \iint z\partial_a\lambda_t(a,z)\, g_t(a,z)\, dadz + \iint z\mu_t g_t(a,z)\, dadz \\ &+ \theta_t\frac{\epsilon_t}{\delta}\iint z\frac{d\tau_t^L(a,z)}{dc_t(a,z)} g_t(a,z)\, dadz - \iint z\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz - \underline{z}\xi_t^{HTM} g_t(\underline{a},\underline{z}), \end{aligned}$$

where $\iint z\mu_t g_t(a,z)\, da\, dz = \mu_t$.

Combining both of these equations, we obtain

$$\begin{aligned} 0 = &\iint \left(zu'(c_t(a,z)) - \frac{v'(n_t(a,z))}{A_t}\right)\left(1 + \frac{\phi_t(a,z)}{g_t(a,z)}\right) g_t(a,z)\, dadz \\ &+ \theta_t\frac{\epsilon_t}{\delta}\iint z\left(\frac{d\tau_t^L(a,z)}{dc_t(a,z)} + \frac{1}{A_t}\frac{d\tau_t^L(a,z)}{dn_t(a,z)}\right) g_t(a,z)\, dadz - \iint z\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz. \end{aligned}$$

Solving for θ_{ss} after imposing that all variables have reached the steady state immediately recovers equation (46) in the text.

B.9 Proof of Proposition 8

The stationary version of the promise-keeping Kolmogorov forward equation (47) follows trivially from the time-varying equation (30), setting $\partial_t \phi_t(a,z) = 0$ and evaluating the RHS at the stationary Ramsey plan.

B.10 Proof of Proposition 9

The optimality conditions of the Ramsey problem imply that

$$\iint a\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz = \iint au'(c_t(a,z))\, g_t(a,z)\, dadz + \theta_t \frac{\epsilon_t}{\delta} \iint a \frac{d\tau_t^L(a,z)}{dc_t(a,z)} g_t(a,z)\, dadz$$
$$- \iint a\partial_a \lambda_t(a,z)\, g_t(a,z)\, dadz - \underline{a}\zeta_t^{HTM} g_t(\underline{a},\underline{z})$$

and

$$\iint z\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz = \mu_t + \iint zu'(c_t(a,z))\, g_t(a,z)\, dadz + \theta_t \frac{\epsilon_t}{\delta} \iint z \frac{d\tau_t^L(a,z)}{dc_t(a,z)} g_t(a,z)\, dadz$$
$$- \iint z\partial_a \lambda_t(a,z)\, g_t(a,z)\, dadz - \underline{z}\zeta_t^{HTM} g_t(\underline{a},\underline{z})$$

Combining these equations with FOCs (33) and (35), we can write, respectively

$$\iint z\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz = \iint \left(zu'(c_t(a,z)) - \frac{v'(n_t(a,z))}{A_t}\right)\left(1 + \frac{\phi_t(a,z)}{g_t(a,z)}\right) g_t(a,z)\, dadz$$
$$+ \theta_t \frac{\epsilon_t}{\delta} \iint z \left(\frac{d\tau_t(a,z)}{dc_t(a,z)} + \frac{1}{A_t}\frac{d\tau_t(a,z)}{dn_t(a,z)}\right) g_t(a,z)\, dadz$$

and

$$\iint a\tilde{\chi}_t(a,z)\, g_t(a,z)\, dadz = \iint a\left(1 + \frac{\phi_t(a,z)}{g_t(a,z)}\right) u'(c_t(a,z))\, g_t(a,z)\, dadz$$
$$+ \theta_t \frac{\epsilon_t}{\delta} \iint a \frac{d\tau_t(a,z)}{dc_t(a,z)} g_t(a,z)\, dadz$$

Finally, expressing $\tilde{\chi}_t(a,z)$ as $\tilde{\chi}_t(a,z) = (1 - \mathcal{M}_t(a,z))\left(\mu_t + \theta_t \frac{\epsilon_t}{\delta}\frac{d\tau_t(a,z)}{dc_t(a,z)}\right)$, we can write

$$\mu_t = \frac{\iint \left(zu'(c_t(a,z)) - \frac{v'(n_t(a,z))}{A_t}\right)\left(1 + \frac{\phi_t(a,z)}{g_t(a,z)}\right) g_t(a,z)\, dadz + \theta_t \frac{\epsilon_t}{\delta} \iint z \left(\frac{1}{A_t}\frac{d\tau_t(a,z)}{dn_t(a,z)} + \mathcal{M}_t(a,z)\frac{d\tau_t(a,z)}{dc_t(a,z)}\right) g_t(a,z)\, dadz}{\iint z(1 - \mathcal{M}_t(a,z))\, g_t(a,z)\, dadz}$$

and

$$\mu_t = \frac{\iint a\left(1+\frac{\phi_t(a,z)}{g_t(a,z)}\right)u'\left(c_t\left(a,z\right)\right)g_t\left(a,z\right)dadz + \theta_t\frac{\epsilon_t}{\delta}\iint a\mathcal{M}_t\left(a,z\right)\frac{d\tau_t(a,z)}{dc_t(a,z)}g_t\left(a,z\right)dadz}{\iint a\left(1-\mathcal{M}_t\left(a,z\right)\right)g_t\left(a,z\right)dadz}$$

Equation (49) follows from combining these two conditions after collecting terms. The logic behind these perturbations is identical to the discretion case, explained in detail in the Proof of Proposition 2.

Output gap targeting rule under isoelastic preferences. Under isoelastic preferences, the targeting rule admits the alternative representation

$$\begin{aligned} Y_t = \tilde{Y}_t \times \left(\frac{\epsilon_t}{\epsilon_t - 1}\frac{1}{1+\tau^L}\right)^{\frac{1}{\gamma+\eta}} \times \Bigg\{ 1 - \Omega_t^D \frac{\iint au'(c_t(a,z))g_t(a,z)\,da\,dz}{\iint zu'(c_t(a,z))g_t(a,z)\,da\,dz} \\ + \frac{\iint (z-\Omega_t^D)u'(c_t(a,z))\phi_t(a,z)\,da\,dz}{\iint zu'(c_t(a,z))g_t(a,z)\,da\,dz} \\ + \theta_t\frac{\epsilon_t}{\delta}\frac{\iint\left(\frac{z}{A_t}\frac{d\tau_t(a,z)}{dN_t} + (z-\Omega_t^D a)\mathcal{M}_t(a,z)\frac{d\tau_t(a,z)}{dc_t(a,z)}\right)g_t(a,z)\,da\,dz}{\iint zu'(c_t(a,z))g_t(a,z)\,da\,dz}\Bigg\}^{\frac{1}{\gamma+\eta}} \end{aligned} \tag{65}$$

in terms of output gaps, where $\tilde{Y}_t$ is natural output in HANK.

C Appendix for Section 3

C.1 Proof of Proposition 1

Having introduced the discretization of the implementability conditions and the Ramsey problem, it is now straightforward to derive the optimality conditions for policy under discretion. We again discretize the planning problem in both individual state variables (a, z) and in time. 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 $n \in \{1, \dots, N\}$. With a step size $dt = \frac{T}{N-1}$, we have $t_n = dt(n-1)$. We assume that a planner under discretion controls policy at time step s and takes as given policy from $s+1$ onwards.

The planning problem under discretion at time s is 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}{dt A_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' \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}{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^w_n)^2 \Bigg] & \\
+ \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}{dt A_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}{dt A_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] & \\
+ \theta_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 sum (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_s^w, N_s, i_s} \sum_{n=s}^{N-1} e^{-\rho t_n} \Bigg\{ \Bigg\{ \\
&+ 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}' \boldsymbol{g}_t - v(N_n) \mathbf{1}' \boldsymbol{g}_t - \frac{\delta}{2} (\pi_n^w)^2 \mathbf{1}' \boldsymbol{g}_t \\
&+ \boldsymbol{\phi}_n' \Bigg[-\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 \Bigg] \\
&+ \boldsymbol{\phi}_n' \boldsymbol{A}^z \boldsymbol{V}_n + \sum_{i \geq 2} \phi_{i,n} \left(i_n a_i - \pi_n^w a_i + \frac{A_{n+1} - A_n}{dt A_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_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 \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] \\
&+ \theta_n \left[-\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 \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}
$$

Crucially, the Markov planner at time step s does realize that her policy decisions affect the evolution of state variables, i.e., the distribution $\boldsymbol{g}_{s+1}$ that the "future planner" at time step $s+1$ takes as her initial condition. She does not internalize, however, that her policy decisions also determine the terminal conditions on forward-looking equations, i.e., inflation and lifetime values, that "past planners" take as given.

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 $\boldsymbol{V}_{s+1}$. The planner under discretion takes the future value $\boldsymbol{V}_{s+1}$ as given. And the planner is constrained by the competitive equilibrium condition that households make consumption decisions *purely* in terms of $\boldsymbol{V}_{s+1}$. By the household's first-order condition, then, $\boldsymbol{c}_s$ is pinned down as a function of $\boldsymbol{V}_{s+1}$.

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

$$\begin{aligned} 0 =& u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}_{s+1}' - e^{\rho dt}\boldsymbol{\lambda}_s'}{dt} + (\boldsymbol{\lambda}_{s+1}'(\boldsymbol{A}^z)')' \\ &+ \frac{d}{d\boldsymbol{g}_{s+1}}\left[\frac{1}{da}(\boldsymbol{D}_a\boldsymbol{\lambda}_{s+1})'[\boldsymbol{s}_{s+1}\cdot\boldsymbol{g}_{s+1}]\right] + \theta_{s+1}N_{s+1}\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_{s+1}\boldsymbol{z}\cdot u'(\boldsymbol{c}_{s+1}). \end{aligned}$$

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

$$0 = u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}'_{s+1} - e^{\rho dt}\boldsymbol{\lambda}'_s}{dt}$$
$$+ \boldsymbol{A}\boldsymbol{\lambda}_{s+1} + \theta_{s+1}N_{s+1}\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_{s+1}\boldsymbol{z}\cdot u'(\boldsymbol{c}_{s+1}).$$

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

$$0 = u'(c_{s,i})g_{s,i} + \mu_s g_{s,i} + u'(c_{s,i})\phi_{s,i} + u''(c_{s,i})\chi_{s,i} + \theta_s N_s \frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s z_i u''(c_{s,i})g_{s,i}$$
$$- \frac{1}{da}\phi_{s,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{V}_s - \frac{1}{da}g_{s,i}\boldsymbol{D}_{a,[i,:]}\boldsymbol{\lambda}_s.$$

Derivative π^w_n. We have

$$0 = \Big[-u'(c_{s,1})g_{s,1}a_1 - \mu_s g_{s,1}a_1 - \delta\pi^w_s\mathbf{1}'\boldsymbol{g}_s - \phi_{s,1}u'(c_{s,1})a_1 - \delta\pi^w_s\boldsymbol{\phi}'_s\mathbf{1}\Big]d\boldsymbol{x}$$
$$- \theta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}a_1 d\boldsymbol{x}$$
$$+ \Big[-\sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}\lambda_{s,i}\frac{\boldsymbol{D}'_{a,[i,:]}}{da}\Big[\begin{pmatrix} 0 \\ -\boldsymbol{a}_{[2:J]}\end{pmatrix}\cdot\boldsymbol{g}_s\Big]d\boldsymbol{x}$$
$$+ \frac{1}{dt}\theta_s.$$

Thus, we have

$$0 = \Big[-u'(c_{s,1})g_{s,1}a_1 - \mu_s g_{s,1}a_1 - \delta\pi^w_s\mathbf{1}'\boldsymbol{g}_s - \phi_{s,1}u'(c_{s,1})a_1 - \delta\pi^w_s\boldsymbol{\phi}'_s\mathbf{1}\Big]d\boldsymbol{x}$$
$$- \theta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s N_s z_1 u''(c_{s,1})g_{s,1}a_1 d\boldsymbol{x}$$
$$- \sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s d\boldsymbol{x} - \sum_{i\geq 2}g_{s,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s d\boldsymbol{x} + \frac{1}{dt}\theta_s.$$

Derivative i_n. The nominal interest rate derivative is parallel to that for 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

$$\begin{aligned}0 =& u'(c_{s,1})g_{s,1}a_1 + \mu_s g_{s,1}a_1 + \phi_{s,1}u'(c_{s,1})a_1 + \sum_{i\geq 2}\phi_{i,s}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}g_{s,i}a_i\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s \\ &+ \theta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_sN_sz_1u''(c_{s,1})g_{s,1}a_1.\end{aligned}$$

Derivative N_n. Finally, we take the derivative for aggregate labor. This is again a static problem. We have

$$\begin{aligned}0 =& \left[u'(c_{s,1})g_{s,1}z_1A_s + \mu_s g_{s,1}z_1A_s + \phi_{s,1}u'(c_{s,1})z_1A_s + \sum_{i\geq 2}\phi_{i,s}z_iA_s\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{V}_s + \sum_{i\geq 2}g_{s,i}z_iA_s\frac{\boldsymbol{D}_{a,[i,:]}}{da}\boldsymbol{\lambda}_s\right]d\boldsymbol{x} \\ &+ \theta_s\frac{\epsilon}{\delta}\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_sN_sz_1u''(c_{s,1})g_{s,1}z_1A_sd\boldsymbol{x} \\ &- v'(N_s)\mathbf{1}'\boldsymbol{g}_sd\boldsymbol{x} - v'(N_s)\boldsymbol{\phi}_s'\mathbf{1}d\boldsymbol{x} \\ &- \mu_sA_s + \theta_s\frac{\epsilon}{\delta}\left(\frac{\epsilon-1}{\epsilon}(1+\tau^L)A_s(\boldsymbol{z}\cdot u'(\boldsymbol{c}_s))'\boldsymbol{g}_sd\boldsymbol{x} - v'(N_s)\right) - \frac{\epsilon}{\delta}\theta_sv''(N_s)N_s.\end{aligned}$$

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

$$\theta_s = 0$$
$$\boldsymbol{\phi}_s = 0$$

because the planner does not respect 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. Using $\theta_s = 0$, we have

$$0 = u(\boldsymbol{c}_{s+1}) + \mu_{s+1}\boldsymbol{c}_{s+1} - v(N_{s+1})\mathbf{1} - \frac{\delta}{2}(\pi^w_{s+1})^2\mathbf{1} + \frac{\boldsymbol{\lambda}'_{s+1} - e^{\rho dt}\boldsymbol{\lambda}'_s}{dt} + \boldsymbol{A}\boldsymbol{\lambda}_{s+1}.$$

The optimality condition for consumption becomes

$$\tilde{\chi}_s = u'(c_s) + \mu_s - \lambda_{a,s},$$

where $\tilde{\chi}_s = -\chi_s \frac{u''(c_s)}{g_s}$. The optimality condition for monetary policy now becomes

$$0 = \Big[u'(c_{s,1}) + \mu_s\Big] g_{s,1} a_1 + \sum_{i \geq 2} g_{s,i} a_i \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{\lambda}_s$$

And finally, the optimality condition for aggregate economic activity becomes

$$0 = \Big[u'(c_{s,1}) g_{s,1} z_1 A_s + \mu_s g_{s,1} z_1 A_s + \sum_{i \geq 2} g_{s,i} z_i A_s \frac{\boldsymbol{D}_{a,[i,:]}}{da} \boldsymbol{\lambda}_s\Big] d\boldsymbol{x} - v'(N_s) - \mu_s A_s.$$

C.2 Proof of Proposition 2

Proving Proposition 2 amounts to a judicious combination of the optimality conditions for policy under discretion (Proposition 1). In particular, we introduce two useful policy perturbations that are purposefully designed so that they have a neutral impact on aggregate excess demand. The first perturbation combines equations (20) and (21), 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

$$\underbrace{\iint \left(z u'(c_t(a,z)) - \frac{v'(N_t)}{A_t} \right) g_t(a,z)\, da\, dz}_{\text{Aggregate Labor Wedge}} - \iint z \tilde{\chi}_t(a,z) g_t(a,z)\, da\, dz = 0.$$

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.[51]

The second perturbation combines equations (20) and (22) 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

[51] 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 (20), we obtain a condition that defines μ_t as a weighted sum of future labor wedges:

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

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.

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 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. Formally, we use the fact that we can write $\chi_t(a,z)$ as

$$\tilde{\chi}_t\,(a,z) = (1 - \mathcal{M}_t\,(a,z))\,\mu_t. \tag{66}$$

Hence, by substituting for $\chi_t(a,z)$ in both the aggregate activity condition and the interest rate condition, and equalizing μ_t, we recover equation (25) in the text. This targeting rule shows that, under discretion, a utilitarian planner in a heterogeneous-agent environment trades off aggregate stabilization against redistribution.

Equation (25) 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.[52] Equation (25) implies that, at an optimum, the planner sets policy trading off aggregate stabilization and redistribution motives.

Derivation of output gap targeting rule for isoelastic preferences. Under isoelastic (CRRA) preferences, we can represent the targeting rule of Proposition 2 in terms of output gaps. We use

$$\iint \frac{zu'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz - Y_t^{\gamma+\eta}A_t^{-(1+\eta)} = \Omega_t^{\mathrm{D}} \iint \frac{au'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz.$$

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

Using natural output definitions

$$\tilde{Y}_t^{\text{RA}} = \left(\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)A_t^{1+\eta}\right)^{\frac{1}{\gamma+\eta}}$$

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

we rearrange and obtain

$$\frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)Y_t^{\gamma+\eta} = \frac{\epsilon_t - 1}{\epsilon_t}(1+\tau^L)A_t^{1+\eta}\left(\iint \frac{zu'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz - \Omega_t^{\text{D}}\iint \frac{au'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz\right)$$

or simply

$$Y_t = \tilde{Y}_t \times \underbrace{\left(\frac{\epsilon_t}{\epsilon_t - 1}\frac{1}{1+\tau^L}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Cost-Push Wedge}} \times \underbrace{\left(1 - \Omega_t^{\text{D}}\frac{\iint \frac{au'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz}{\iint \frac{zu'(c_t(a,z))}{u'(C_t)}g_t(a,z)\,da\,dz}\right)^{\frac{1}{\gamma+\eta}}}_{\text{Redistribution}}$$

C.3 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 (6). Setting $\dot{\pi}_{\text{ss}}^w = 0$, this yields

$$\pi_{\text{ss}}^w = -\frac{\epsilon}{\delta}A_{\text{ss}}N_{\text{ss}}\iint\left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L)zu'(c_{\text{ss}}(a,z)) - \frac{v'(N_{\text{ss}})}{A_{\text{ss}}}\right)g_{\text{ss}}(a,z)\,da\,dz.$$

We now separate terms into a markup component and a redistribution component, analogous to Proposition 2,

$$\pi_{\text{ss}}^w = -\frac{\epsilon}{\delta}A_{\text{ss}}N_{\text{ss}}\iint\left(\frac{\epsilon - 1}{\epsilon}(1+\tau^L)zu'(c_{\text{ss}}(a,z)) - zu'(c_{\text{ss}}(a,z)) + \Omega_{\text{ss}}^{\text{D}}au'(c_{\text{ss}}(a,z))\right)g_{\text{ss}}(a,z)\,da\,dz$$

or simply

$$\pi_{\text{ss}}^w = \frac{\epsilon}{\delta}A_{\text{ss}}N_{\text{ss}}\left[\left(1 - \frac{\epsilon - 1}{\epsilon}(1+\tau^L)\right)\iint zu'(c_{\text{ss}}(a,z))g_{\text{ss}}(a,z)\,da\,dz - \Omega_{\text{ss}}^{\text{D}}\iint au'(c_{\text{ss}}(a,z))g_{\text{ss}}(a,z)\,da\,dz\right],$$

which concludes the proof.

D 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. (2022).[53] 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 $c_{t_n}(a_i, z_i)$.

D.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 $(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z})$ to time paths of aggregate outputs $\boldsymbol{Y}$. We use bold-faced notation here to indicate time paths, with $\boldsymbol{i} = \{i_n\}_{n=1}^N$. It will be useful to explicitly distinguish between the time paths for policy $\boldsymbol{i}$ 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

[53] For a detailed description of the discretization procedure, see Achdou et al. (2022) or Schaab and Zhang (2022). We also leverage the adaptive sparse grid method developed by Schaab and Zhang (2022) and Schaab (2020) to solve dynamic programming problems in continuous time.

an individual outcome.[54] For example, aggregate consumption takes the form $C_n = \boldsymbol{c}_n' \boldsymbol{g}$. Given an initial distribution $\boldsymbol{g}_0$, aggregate outcomes $\boldsymbol{Y}$ then solve the system of equations

$$\boldsymbol{V}_n = v(\boldsymbol{V}_{n+1}, X_n, i_n, Z_n) \tag{67}$$

$$\boldsymbol{g}_{n+1} = \Lambda(\boldsymbol{V}_{n+1}, X_n, i_n, Z_n)\boldsymbol{g}_n \tag{68}$$

$$Y_n = y(\boldsymbol{V}_{n+1}, X_n, i_n, Z_n)' \boldsymbol{g}_n. \tag{69}$$

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 $\boldsymbol{i}$, and shocks $\boldsymbol{Z}$. In summary, equilibria given policy and shocks can be expressed in terms of the *equilibrium map* $\mathcal{H}(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z}) = 0$.

D.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 (67) through (69). 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 (52) 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 $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i})$ and $\boldsymbol{Z}$.

Forward-looking individual multipliers take the form

$$\boldsymbol{\lambda}_n = f(\boldsymbol{\lambda}_{n+1}, \boldsymbol{V}_n, X_n, M_n, i_n, Z_n), \tag{70}$$

which is analogous to equation (67), which characterizes individual forward-looking behavior. For example, it is straightforward to verify that equation (31) 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 (67), the recursive structure of forward-looking individual multipliers

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

allows us to efficiently compute their first-order derivatives. We summarize this observation in the following Lemma.

Lemma 22. *For any $k \geq 1$, we have*

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

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

Lemma 22 represents an extension of Auclert et al. (2021)'s fake-news result to the multipliers that emerge from Ramsey problems. The unifying insight here is that, just like competitive equilibrium in heterogeneous-agent economies comprises a forward-backward system of dynamic equations, so do Ramsey plans. In other words, equation (31), which defines the social lifetime value $\lambda_t(a, z)$ still satisfies a (Hamilton-Jacobi-) Bellman equation. The "fake-news property" identified by Auclert et al. (2021) applies to any backward equation, including those satisfied by multipliers.

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 (30). Consequently, backward-looking multipliers can be represented as

$$\boldsymbol{\phi}_{n+1} = \Lambda(\boldsymbol{V}_{n+1}, X_n, i_n, Z_n)\boldsymbol{\phi}_n + b(\boldsymbol{\phi}_n, \boldsymbol{\lambda}_n, \boldsymbol{V}_n, \boldsymbol{g}_n, X_n, M_n, i_n, Z_n). \tag{71}$$

The same arguments developed by Auclert et al. (2021) for efficiently computing sequence-space derivatives of the cross-sectional distribution also apply to multipliers that satisfy (Kolmogorov) forward equations.

The multiplier representations (70) and (71), together with equations (67) through (69) and the equilibrium map (51), let us conclude that timeless Ramsey plans admit the sequence-space representation (52).

D.3 Sequence-Space Perturbations in the Primal: Proof of Proposition 11

Denote the Ramsey plan by $\boldsymbol{R} = (\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{i})$. 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 (52).

We now show that $\mathcal{R}(\boldsymbol{R}_{\text{ss}},\boldsymbol{Z}_{\text{ss}})=0$ as well, assuming, as we do in Proposition 11, that the Ramsey problem is initialized at $(g_0,\phi,\theta)=(g_{\text{ss}},\phi_{\text{ss}},\theta_{\text{ss}})$. 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 $\frac{d}{d\boldsymbol{i}}L^{\text{TP}}(\boldsymbol{\theta}_{\text{ss}},\boldsymbol{Z}_{\text{ss}})=0$ also implies that each of the associated first-order conditions of the timeless primal problem are 0 when evaluated at $(\boldsymbol{R}_{\text{ss}},\boldsymbol{Z}_{\text{ss}})$. Putting these observations together implies $\mathcal{R}(\boldsymbol{R}_{\text{ss}},\boldsymbol{Z}_{\text{ss}})=0$.

We are then simply left with

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

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

D.4 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 (52) but a sequence-space representation of the timeless dual Lagrangian, which we now introduce.

The timeless dual Lagrangian defined in Appendix B.6 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 (51) 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

$$L^{\text{TD}}(\boldsymbol{X}(\boldsymbol{i},\boldsymbol{Z}),\boldsymbol{i},\boldsymbol{Z}),\tag{72}$$

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

The timeless dual Lagrangian (72) implies the local efficiency criterion for optimal policy

$$\mathcal{F}(\boldsymbol{i},\boldsymbol{Z})=\frac{d}{d\boldsymbol{i}}L^{\text{TD}}(\boldsymbol{X}(\boldsymbol{i},\boldsymbol{Z}),\boldsymbol{i},\boldsymbol{Z})=0,\tag{73}$$

where $\mathcal{F}(\cdot)$ implicitly takes as given an initial distribution as well as initial promises. Equation (73) 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 (51). It only influences the planner's assessment of optimal policy.

Proposition 23. (**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 $g_0(a,z) = g_{\text{ss}}(a,z)$, 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 by*

$$d\boldsymbol{i} = -\mathcal{F}_{\boldsymbol{i}}^{-1}\mathcal{F}_{\boldsymbol{Z}} d\boldsymbol{Z}, \tag{74}$$

where $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{\text{ss}}$ 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 23 at the end of this subsection.

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 D.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.

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 (52) 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 (52), multipliers $\boldsymbol{M}$ enter as an explicit argument. It is not generically possible to characterize Ramsey plans in the spirit of (52) 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.

Proof of Proposition 23. We proceed as follows. A first-order Taylor approximation of $F(\cdot)$ (in $\boldsymbol{i}$ and $\boldsymbol{Z}$) around the stationary Ramsey plan yields

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

First, we must have

$$F(\boldsymbol{g}_{\text{ss}}, \boldsymbol{\phi}_{\text{ss}}, \theta_{\text{ss}}, \boldsymbol{i}, \boldsymbol{Z}) = 0$$

by construction because this is our definition for optimal policy $\boldsymbol{i}(\boldsymbol{Z})$. Second, we also have

$$F(\boldsymbol{g}_{\text{ss}}, \boldsymbol{\phi}_{\text{ss}}, \theta_{\text{ss}}, \boldsymbol{i}_{\text{ss}}, \boldsymbol{Z}_{\text{ss}}) = 0,$$

which is the main result of Section 4.3 and whose proof is in Appendix B.7. Denoting $d\boldsymbol{i} = \boldsymbol{i} - \boldsymbol{i}_{\text{ss}}$ and $d\boldsymbol{Z} = \boldsymbol{Z} - \boldsymbol{Z}_{\text{ss}}$, we thus have

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

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

D.5 Sequence-Space Hessians

To compute optimal policy in the dual using Proposition 23, we effectively need to differentiate $L^{\text{TD}}(\cdot)$ twice. In particular, $F(\cdot) = \frac{d}{d\boldsymbol{i}} L^{\text{TD}}$ features first-order derivative terms, which can be cast as sequence-space Jacobians (Auclert et al., 2021). Therefore, computing the total derivatives $\frac{d}{d\boldsymbol{i}} F(\cdot)$ and $\frac{d}{d\boldsymbol{Z}} F(\cdot)$, which are used in equation (74) to characterize optimal policy $d\boldsymbol{i}$, 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.

In a sequence-space representation of our model, these objects are all functions of the time paths of aggregate inputs, i.e., $(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z})$, where $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z})$. 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 $\frac{d}{d\boldsymbol{i}} F$ 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 D.5.1. Subsequently, in Section D.5.2, we show how to build up the second-order total derivatives of the timeless dual Lagrangian, i.e., $\frac{d}{d\boldsymbol{i}} F$ and $\frac{d}{d\boldsymbol{Z}} F$, from sequence-space Hessians.

D.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 $\frac{\partial^2}{\partial i_k \partial i_l} Y_n$ in a sequence-space representation of the model. The results we present below hold for any mixed partial derivative of $Y_n(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z})$, but to ease notation we focus specifically on the mixed derivative $\frac{\partial^2}{\partial i_k \partial i_l}$ for some given $k, l \in \{1, \ldots, N\}$.

Using equation (69), we can rewrite the mixed derivative of aggregate outcome Y_n at time t_n as

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

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 $\frac{\partial^2}{\partial i_k \partial i_l}\boldsymbol{y}_n$ and $\frac{\partial^2}{\partial i_k \partial i_l}\boldsymbol{g}_n$, 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

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

Second, the recursive structure of the system (67) - (69) gives rise to the following key property of mixed partial derivatives in sequence space.

Lemma 24. *We have*

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

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

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial i_k \partial i_N}$$

for $1 \leq k \leq N$, which requires only N numerical derivative evaluations.[55]

Third, we exploit the fact that the transition matrix Λ, which describes the law of motion of the cross-sectional distribution in equation (69), has a particular structure in continuous time. In particular, we have

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

where $\boldsymbol{A}$ is the $J \times J$ matrix that discretizes the HJB operator $\mathcal{A}$, and $\boldsymbol{A}'$, its transpose, is the analog

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

for the adjoint $\mathcal{A}^*$. In our baseline HANK model, the discretized transition matrix takes the form

$$\boldsymbol{A}_n = \boldsymbol{s}_n \cdot \boldsymbol{D}_a + \boldsymbol{A}^z \tag{75}$$

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, 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 (69) simply amounts to differentiating $\boldsymbol{s}_n$ in equation (75).[56] We record this observation in the following Lemma.

Lemma 25. *The first- and second-order mixed partial derivatives of the transition matrix Λ_n in our setting are given by*

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

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 corresponds to the stationary Ramsey plan, that is, $\boldsymbol{g}_1 = \boldsymbol{g}_{\text{ss}}$, 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, $\frac{\partial^2 \boldsymbol{g}_1}{\partial i_k \partial i_l} = 0$. Using equation (69) and Lemma 25, the response of the cross-sectional distribution at time step $n = 2$ is thus

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

We now exploit the recursive structure of equation (69) to derive two alternative expressions for the mixed derivatives $\frac{\partial^2 \boldsymbol{g}_n}{\partial i_k \partial i_l}$, for $n \geq 3$. We summarize in the next Lemma.

Lemma 26. *The mixed partial derivatives of the cross-sectional distribution $\boldsymbol{g}_n$ at time steps $n \geq 3$ can be computed recursively using*

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

[56] Notice that equation (75) is specific to our baseline model and consequently breaks with the spirit of generality otherwise adopted in this section. However, an equation *like* (75) 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 (75) in the following while otherwise maintaining our general notation.

or non-recursively using

$$\frac{\partial^2 \boldsymbol{g}_n}{\partial i_k \partial i_l} = \sum_{r=1}^{R_1} (\Lambda_{ss})^{n-r-1} \frac{\partial^2 \boldsymbol{g}_2}{\partial i_k \partial i_l} + dt \sum_{r=1}^{R_2} (\Lambda_{ss})^{n-r-2} \left(\frac{\partial \boldsymbol{s}_1}{\partial i_{k-r}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{1+r}}{\partial i_l} + dt \sum_{r=1}^{R_3} (\Lambda_{ss})^{n-r-2} \left(\frac{\partial \boldsymbol{s}_1}{\partial i_{l-r}} \cdot \boldsymbol{D}_a \right) \frac{\partial \boldsymbol{g}_{1+r}}{\partial i_k}$$

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

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial i_k \partial i_l} = \frac{\boldsymbol{y}_n^{++} - \boldsymbol{y}_n^{+-} - \boldsymbol{y}_n^{-+} + \boldsymbol{y}_n^{--}}{4h^2} \tag{76}$$

where $\boldsymbol{y}_n^{++} = \boldsymbol{y}_n(\ldots, i_k + h, \ldots, i_l + h, \ldots)$, $\boldsymbol{y}_n^{+-} = \boldsymbol{y}_n(\ldots, i_k + h, \ldots, i_l - h, \ldots)$, $\boldsymbol{y}_n^{-+} = \boldsymbol{y}_n(\ldots, i_k - h, \ldots, i_l + h, \ldots)$, and $\boldsymbol{y}_n^{--} = \boldsymbol{y}_n(\ldots, i_k - h, \ldots, i_l - h, \ldots)$. 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

$$\frac{\partial^2 \boldsymbol{y}_n}{\partial i_k \partial i_l} = \frac{\boldsymbol{y}_n^{++} - \boldsymbol{y}_n^{+\cdot} - \boldsymbol{y}_n^{\cdot+} + 2\boldsymbol{y}_n - \boldsymbol{y}_n^{-\cdot} - \boldsymbol{y}_n^{\cdot-} + \boldsymbol{y}_n^{--}}{2h^2} \tag{77}$$

where $\boldsymbol{y}_n^{+\cdot} = \boldsymbol{y}_n(\ldots, i_k + h, \ldots, i_l, \ldots)$, $\boldsymbol{y}_n^{\cdot+} = \boldsymbol{y}_n(\ldots, i_k, \ldots, i_l + h, \ldots)$, $\boldsymbol{y}_n^{-\cdot} = \boldsymbol{y}_n(\ldots, i_k - h, \ldots, i_l, \ldots)$, and $\boldsymbol{y}_n^{\cdot-} = \boldsymbol{y}_n(\ldots, i_k, \ldots, i_l - h, \ldots)$. Stencil (77) requires only 2 new function evaluations compared to stencil (76)'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.

D.5.2 Total Derivatives and General Equilibrium

Our perturbation approach to optimal stabilization policy in the dual requires the two total derivatives $\frac{d}{d\boldsymbol{i}}F$ 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

$$(F_{\boldsymbol{i}})_{[k,l]} = \sum_{n=1}^{N} e^{-\rho t_n} \frac{d^2 U_n}{di_k di_l} dt + (\boldsymbol{\phi}_{\text{ss}})' \frac{d^2 \boldsymbol{V}_1}{di_k di_l} + \theta_{\text{ss}} \frac{d^2 \pi_1^w}{di_k di_l}, \tag{78}$$

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 (78) 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 (51), which is a system of N equation, assuming for now that $X_n \in \mathbb{R}$. Given paths for policy $\boldsymbol{i}$ and the exogenous shock $\boldsymbol{Z}$, we can solve equation (51) for $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{i}, \boldsymbol{Z})$.

To compute the total derivative $F_{\boldsymbol{i}}$, 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 $\boldsymbol{X}_{i_k} = -\boldsymbol{H}_{\boldsymbol{X}}^{-1} \boldsymbol{H}_{i_k}$. Likewise, the mixed partial derivatives are given by

$$\boldsymbol{X}_{i_k i_l} = -\boldsymbol{H}_{\boldsymbol{X}}^{-1} \boldsymbol{H}_{i_k i_l} + \boldsymbol{H}_{\boldsymbol{X}}^{-1} \boldsymbol{H}_{\boldsymbol{X} i_l} \boldsymbol{H}_{\boldsymbol{X}}^{-1} \boldsymbol{H}_{i_k}. \tag{79}$$

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, $Y(\boldsymbol{X}, \boldsymbol{i}, \boldsymbol{Z})$.

Lemma 27. *The total derivatives of Y are given by*

$$\frac{d^2Y}{di_k di_l} = \begin{pmatrix} Y_{X_1\boldsymbol{X}}\boldsymbol{X}_{i_l} & \dots & Y_{X_N\boldsymbol{X}}\boldsymbol{X}_{i_l} \end{pmatrix} \boldsymbol{X}_{i_k} + Y_{\boldsymbol{X}i_l}\boldsymbol{X}_{i_k} + Y_{\boldsymbol{X}}\boldsymbol{X}_{i_k i_l} + Y_{i_k\boldsymbol{X}}\boldsymbol{X}_{i_l} + Y_{i_k i_l} \tag{80}$$

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

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.

D.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.

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 (77)
   a: construct policy Hessians ▷ use Lemma 24
   b: construct distribution Hessians ▷ use Lemmas 25 and 26

4: Use Hessians to compute mixed derivatives of $\boldsymbol{H}$ and $\boldsymbol{X}$ ▷ use equations (51) and (79)

5: Compute total derivatives for $F_{\boldsymbol{i}}$ and $F_{\boldsymbol{Z}}$ ▷ use equations (78) and (80)

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

D.5.4 Accuracy and Performance

We test the accuracy of our method in Appendix F.3. We show there 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 $d\boldsymbol{i} = -F_{\boldsymbol{i}}^{-1}F_{\boldsymbol{Z}}d\boldsymbol{Z}$. 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.

E RANK with Wage Rigidity

In this Appendix, we present a self-contained treatment of optimal monetary policy in the RANK limit of our 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.

It is well known that the Ramsey problem in the standard New Keynesian model can be represented with a single implementability condition, the Phillips curve. The goal of this Appendix is to parallel our derivations for HANK and facilitate comparison where possible. We therefore represent the Ramsey problem in terms of the same set of implementability conditions we use in the main text. The following Lemma summarizes these for the RANK limit.

Lemma 28. *The implementability conditions that a Ramsey planner faces in RANK can be summarized as*

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

$$C_t = A_t N_t$$

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

E.1 Standard Ramsey Problem

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

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

Crucially, both C_0 and π^w_0 are *free* from the planner's perspective.

Proposition 29. *The first-order conditions for optimal monetary policy in RANK are given by*

$$0 = C^{-\gamma} - \mu + \phi\frac{1}{\gamma}\left(i - \pi^w + \frac{\dot{A}}{A} - \rho\right) - \rho\phi + \dot{\phi} + \theta\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}Au''(C)N \tag{81}$$

$$0 = -\delta\pi^w - \phi\frac{1}{\gamma}C + \vartheta\rho - \rho\vartheta + \dot{\vartheta} \tag{82}$$

$$0 = -v'(N) + \mu A + \vartheta\frac{\epsilon}{\delta}\left((1+\tau^L)\frac{\epsilon-1}{\epsilon}Au'(C) - v'(N)\right) - \theta\frac{\epsilon}{\delta}v''(N)N \tag{83}$$

$$0 = \phi\frac{1}{\gamma}C, \tag{84}$$

with initial conditions

$$0 = \phi_0$$
$$0 = \vartheta_0.$$

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

$$0 = C^{-\gamma} - \mu + \theta\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}Au''(C)N$$
$$\dot{\theta} = -\delta\pi^w$$
$$0 = -v'(N) + \mu A + \theta\frac{\epsilon}{\delta}(1+\tau^L)\frac{\epsilon-1}{\epsilon}Au'(C) - \theta\frac{\epsilon}{\delta}v'(N) - \theta\frac{\epsilon}{\delta}v''(N)N$$

with initial condition $\theta_0 = 0$. The stationary Ramsey plan satisfies

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

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

Proof of Proposition 29. We now integrate by parts and consider a general functional perturbation, yielding

$$
\begin{aligned}
L = \int_0^\infty e^{-\rho t} \Bigg\{ & \frac{1}{1-\gamma}(C+\alpha h_C)^{1-\gamma} - v(N+\alpha h_N) - \frac{\delta}{2}(\pi^w + \alpha h_\pi)^2 \\
& + \phi \left[\frac{1}{\gamma} \left(i + \alpha h_i - \pi^w - \alpha h_\pi + \frac{\dot{A}}{A} - \rho \right)(C+\alpha h_C) \right] \\
& + \mu \left[A(N+\alpha h_N) - C - \alpha h_C \right] \\
& + \theta \left[\rho(\pi^w + \alpha h_\pi) + \frac{\epsilon}{\delta} \left((1+\tau^L)\frac{\epsilon-1}{\epsilon} A u'(C+\alpha h_C) - v'(N+\alpha h_N) \right)(N+\alpha h_N) \right] \\
& - \rho\phi(C+\alpha h_C) + (C+\alpha h_C)\dot{\phi} - \rho\theta(\pi^w+\alpha h_\pi) + (\pi^w + \alpha h_\pi)\dot{\theta} \Bigg\} dt \\
& + \phi_0(C_0 + \alpha h_{C,0}) + \vartheta_0(\pi^w_0 + \alpha h_{\pi,0})
\end{aligned}
$$

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

$$
\begin{aligned}
0 = \int_0^\infty e^{-\rho t} \Bigg\{ & C^{-\gamma} h_C - v'(N) h_N - \delta \pi^w h_\pi \\
& + \phi \left[\frac{1}{\gamma} \left(h_i - h_\pi \right) C + \frac{1}{\gamma} \left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) h_C \right] \\
& + \mu \left[A h_N - h_C \right] \\
& + \theta \Bigg[\rho h_\pi + \frac{\epsilon}{\delta} \left((1+\tau^L)\frac{\epsilon-1}{\epsilon} A u''(C) h_C - v''(N) h_N \right) N \\
& \quad + \frac{\epsilon}{\delta} \left((1+\tau^L)\frac{\epsilon-1}{\epsilon} A u'(C) - v'(N) \right) h_N \Bigg] \\
& - \rho\phi h_C + h_C \dot{\phi} - \rho\theta h_\pi + h_\pi \dot{\theta} \Bigg\} dt + \phi_0 h_{C,0} + \theta_0 h_{\pi,0}
\end{aligned}
$$

Grouping terms,

$$
\begin{aligned}
0 = \int_0^\infty e^{-\rho t} \Bigg\{ & \left[C^{-\gamma} - \mu + \phi \frac{1}{\gamma} \left(i - \pi^w + \frac{\dot{A}}{A} - \rho \right) - \rho \phi + \dot{\phi} + \theta \frac{\epsilon}{\delta} (1 + \tau^L) \frac{\epsilon - 1}{\epsilon} A u''(C) N \right] h_C \\
& + \left[- \delta \pi^w - \phi \frac{1}{\gamma} C + \theta \rho - \rho \theta + \dot{\vartheta} \right] h_\pi \\
& + \left[- v'(N) + \mu A + \vartheta \frac{\epsilon}{\delta} \left((1 + \tau^L) \frac{\epsilon - 1}{\epsilon} A u'(C) - v'(N) \right) - \theta \frac{\epsilon}{\delta} v''(N) N \right] h_N \\
& + \left[\phi \frac{1}{\gamma} C \right] h_i \Bigg\} dt + \phi_0 h_{C,0} + \theta_0 h_{\pi,0}
\end{aligned}
$$

The fundamental lemma of the calculus of variations yields the desired result. Since C_0 and π_0 are both free, it follows that optimality requires $\phi_0 = \theta_0 = 0$. Finally, it follows directly from $0 = \frac{1}{\gamma} \phi_t C_t$ that we must have $\phi_t = 0$ for all t.

E.2 Timeless Ramsey Problem

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

$$
L^{\mathrm{TD}}(\theta) = \int_0^\infty e^{-\rho t} \left\{ \frac{1}{1-\gamma} C_t^{1-\gamma} - v(N_t) - \frac{\delta}{2} (\pi_t^w)^2 \right\} dt \underbrace{-\theta \pi_0^w}_{\text{Inflation Penalty}}
$$

Lemma 30. *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,*

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

Proof. Suppose we differentiate

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

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

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

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

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

Thus,

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

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

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

Thus, we have

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

Finally, we now have

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

which concludes the proof.

■

Our constructive proof of Lemma 30 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.

E.3 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. 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 leverage the results we derive here in Section 4.6 of the main text to compare optimal policy and targeting rules across RANK and HANK.

Following Galí (2015), we define the *natural level of output*, denoted $\tilde{Y}_t$, as the equilibrium level of output under flexible prices. From the Phillips curve, which is in our setting given by

$$\tilde{Y}_t = \left[(1+\tau^L)\frac{\epsilon_t-1}{\epsilon_t}A_t^{1+\eta}\right]^{\frac{1}{\gamma+\eta}}. \tag{85}$$

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

$$\dot{\pi}_t^w = \rho_t \pi_t^w + \frac{\epsilon_t}{\delta}\left[\tilde{Y}_t^{\gamma+\eta} - Y_t^{\gamma+\eta}\right] Y_t^{1-\gamma} A_t^{-1-\eta} \tag{86}$$

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

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

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 31. (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

$$Y_t = \tilde{Y}_t \left(\frac{\frac{1}{1+\tau^L}\frac{\epsilon_t}{\epsilon_t - 1} + \frac{\epsilon_t}{\delta}\theta_t(1-\gamma)}{1 + \frac{\epsilon_t}{\delta}\theta_t(1+\eta)}\right)^{\frac{1}{\gamma+\eta}} \tag{87}$$

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

$$Y_t = \tilde{Y}_t \left(\frac{1 + \frac{\epsilon}{\delta}\theta_t(1-\gamma)}{1 + \frac{\epsilon}{\delta}\theta_t(1+\eta)}\right)^{\frac{1}{\gamma+\eta}}. \tag{88}$$

A solution to the non-linear Ramsey plan is then given by $Y_t = \tilde{Y}_t$, $\theta_t = 0$, 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

$$
\begin{aligned}
L = \int_0^\infty e^{-\int_0^t \rho_s ds} \Bigg\{ & u(Y_t) - v\left(\frac{Y_t}{A_t}\right) - \frac{\delta}{2}(\pi_t^w)^2 \\
& + \theta_t \left[\rho_t \pi^w + \frac{\epsilon_t}{\delta}\left((1+\tau^L)\frac{\epsilon_t - 1}{\epsilon_t} A_t u'(Y_t) - v'\left(\frac{Y_t}{A_t}\right)\right)\frac{Y_t}{A_t} \right] \\
& - \rho_t \theta_t \pi_t^w + \pi_t^w \dot{\theta}_t \Bigg] \Bigg\} dt + \theta_0 \pi_0^w
\end{aligned}
$$

The two first-order conditions are then given by

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

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

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

Further rearranging yields

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

or simply

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

Using the definition of natural output, we therefore have

$$
\left[1 + \frac{\epsilon_t}{\delta}\theta_t(1+\eta)\right]^{\frac{1}{\gamma+\eta}} Y_t = \left(\frac{1}{1+\tau^L}\frac{\epsilon_t}{\epsilon_t - 1} + \frac{\epsilon_t}{\delta}\theta_t (1-\gamma)\right)^{\frac{1}{\gamma+\eta}} \tilde{Y}_t
$$

which concludes the proof.

■

Importantly, the targeting rule of Proposition 31 echoes the 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.

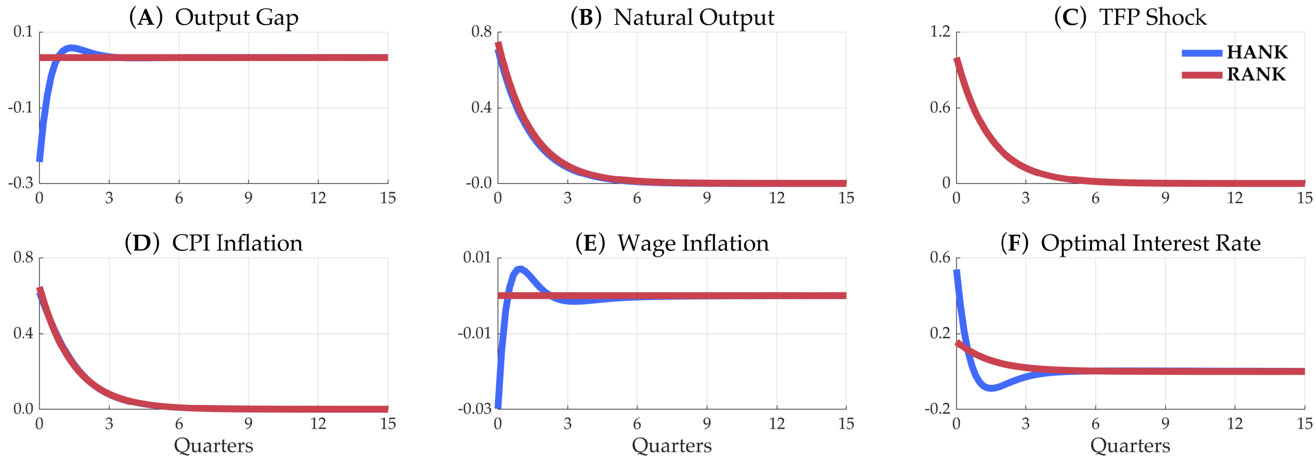


Figure 5. Optimal Policy Transition Dynamics: TFP Shock

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.

F Quantitative Analysis: Additional Results and Robustness

F.1 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.4 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. 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 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

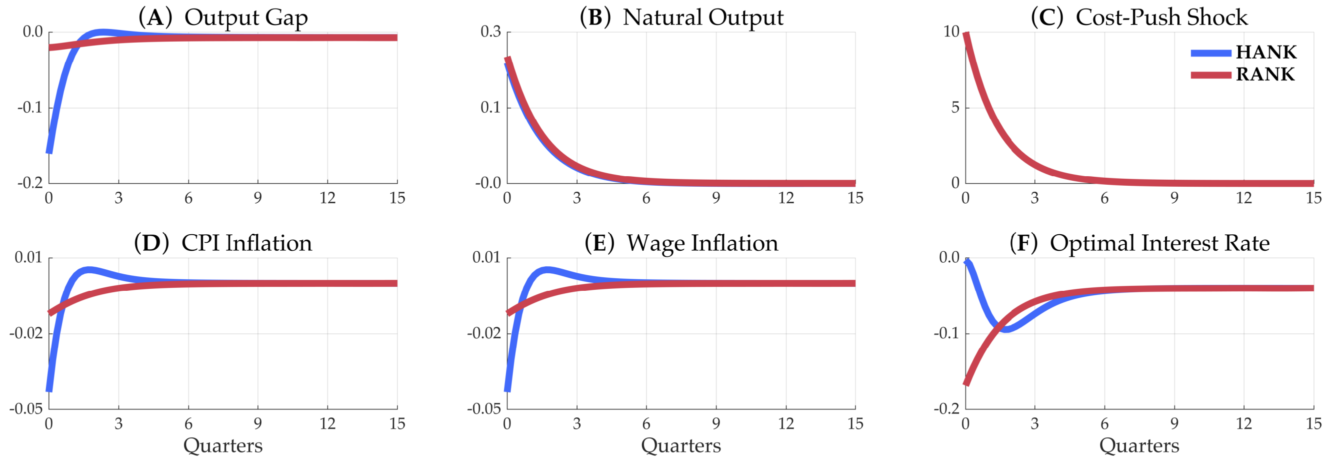


Figure 6. Optimal Policy Transition Dynamics: Cost-Push Shock

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, $\frac{Y_t - \tilde{Y}_t}{\tilde{Y}_t}$, 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.

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.

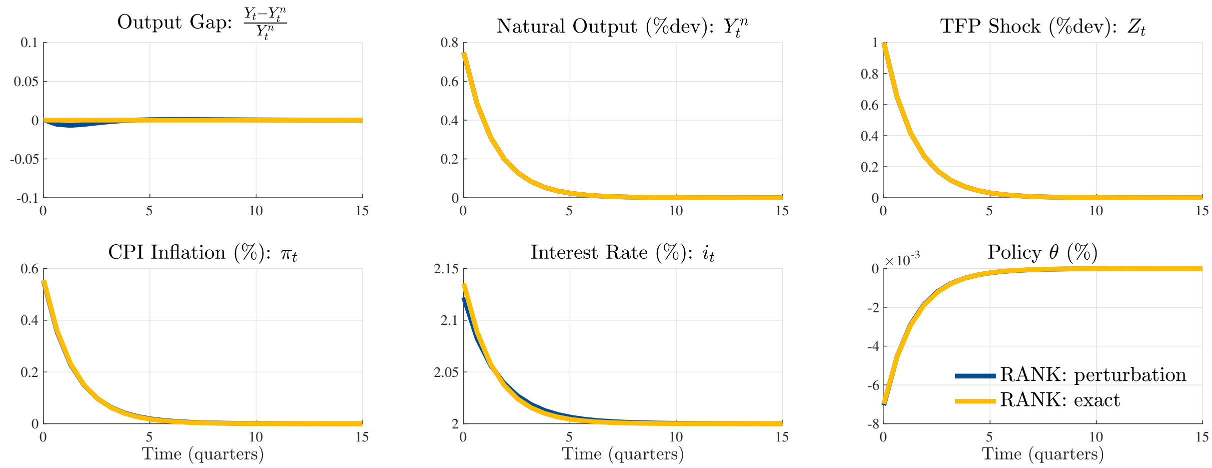
F.2 Cost-Push Shocks

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.

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.

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).

F.3 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}.

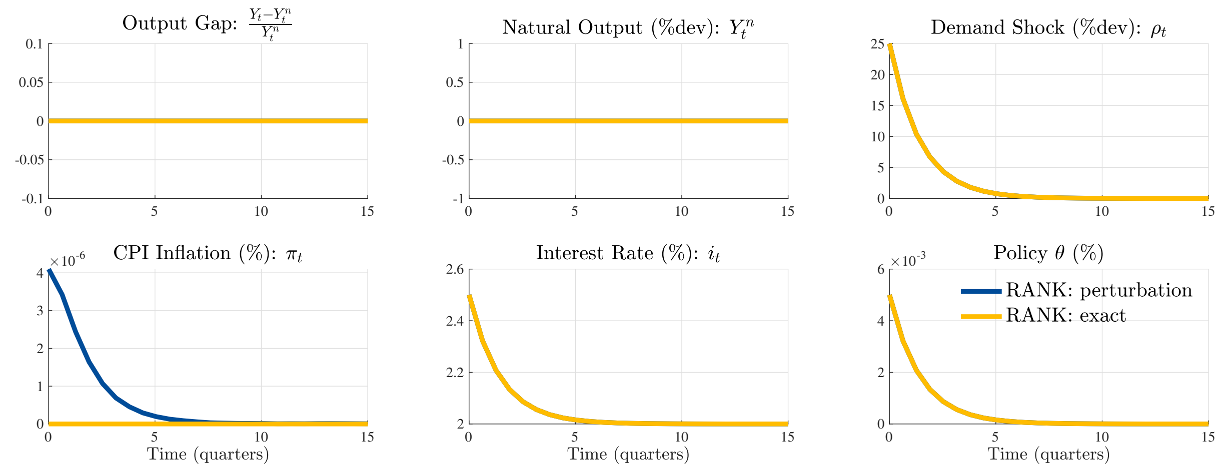
F.4 Transition Dynamics without Optimal Policy

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

$$i_t = r^{ss} + \lambda_\pi \pi_t, \tag{89}$$

where we calibrate $\lambda_\pi = 1.5$.

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).

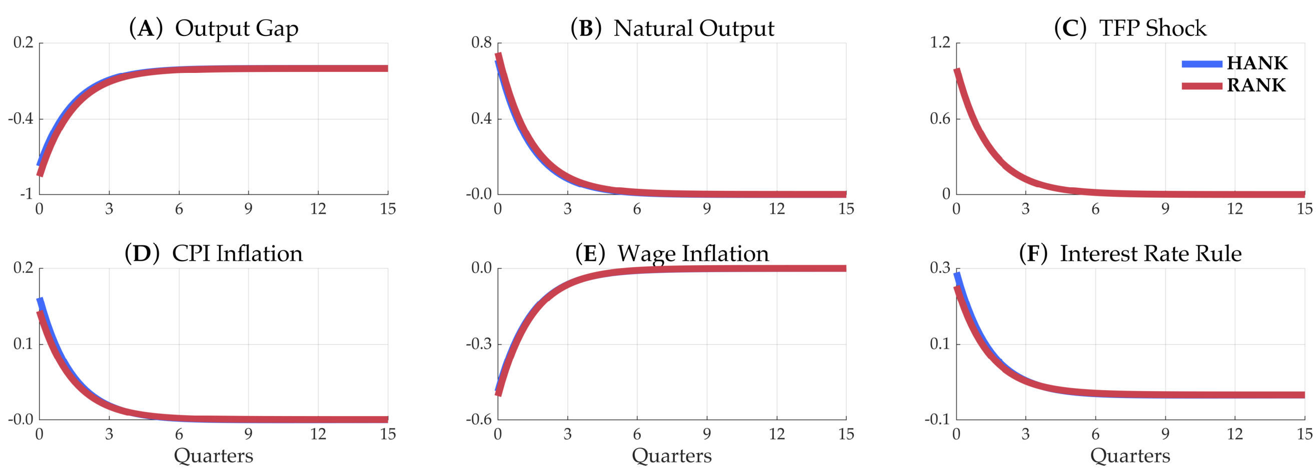


Figure 9. Transition Dynamics under Taylor Rule: TFP Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (89) 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.

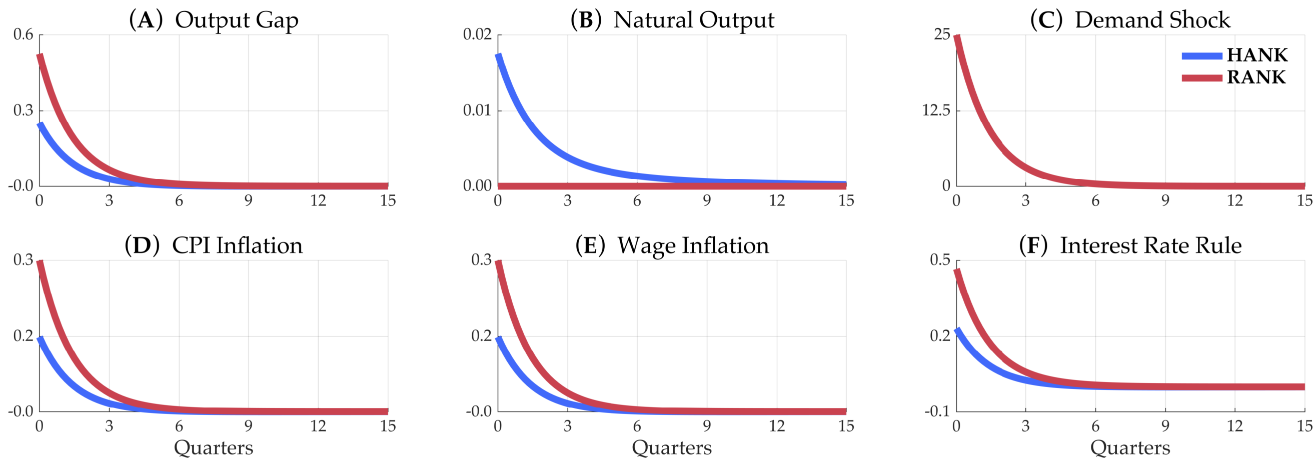


Figure 10. Transition Dynamics under Taylor Rule: Demand Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (89) 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.

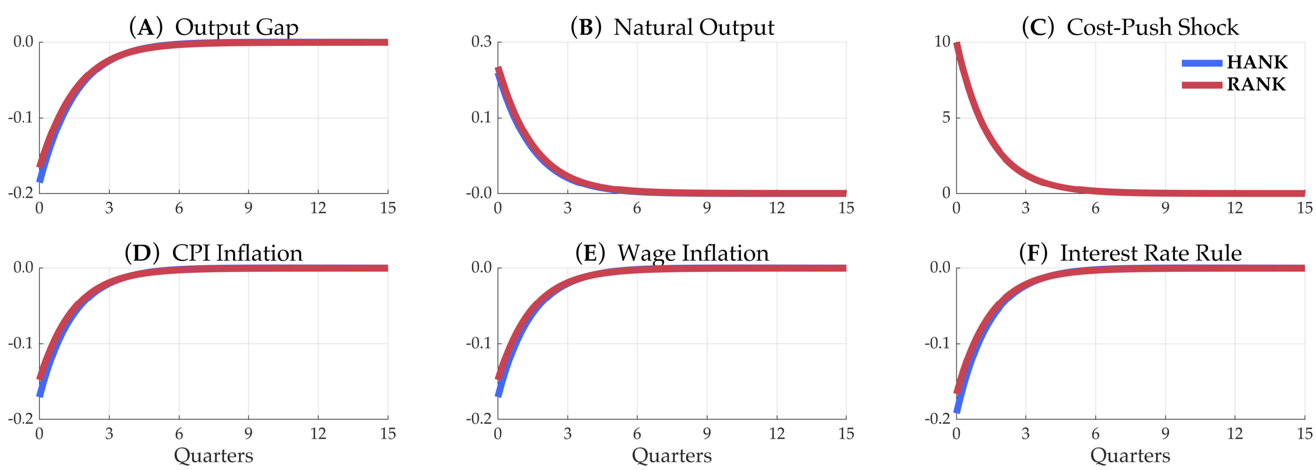


Figure 11. Transition Dynamics under Taylor Rule: Cost-Push Shock

Note. Impulse responses to positive cost-push shock in both RANK (yellow) and HANK (blue) models. The nominal interest rate follows the Taylor rule (89) 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.