Should monetary policy care about redistribution? Optimal monetary and fiscal policy with heterogeneous agents

François Le Grand Alaïs Martin-Baillon Xavier Ragot*

December 18, 2022

Abstract

We derive optimal monetary policy in a heterogeneous-agent economy with nominal frictions and aggregate shocks, considering three assumptions about fiscal policy. First, when time-varying linear taxes on capital and labor can be optimally set, we theoretically prove that optimal monetary policy implements price stability. This implies that monetary policy should focus on ensuring price stability and let fiscal policy deal with redistribution. Second, using both a sequence of simple models and a quantitatively relevant setup, we show that under a standard calibration, the optimal inflation volatility remains low – but positive – when tax rates are constant, and that it tends towards zero when we allow for simple time-varying exogenous tax rates. Third, we consider a constrained-optimal fiscal policy, in which we fix some fiscal tools and let the others be optimally set. We then find that the optimal inflation volatility is also close to zero. In all three cases we find that fiscal policy is more efficient than monetary policy to provide insurance against aggregate shocks.

Keywords: Heterogeneous agents, optimal Ramsey policies, monetary policy, fiscal policy.

JEL codes: D31, E52, D52, E21.

*We thank Tom Sargent for stimulating discussion. We thank Édouard Challe, Albert Marcet, Kurt Mitman, Christian Moser, Galo Nuño, Morten Ravn, Anna Rogantini, Carlos Thomas, and Gianluca Violante for helpful discussions and remarks. We would like to thank seminar participants at the T2M Paris-Dauphine Conference, at the 3rd Conference on Financial Markets and Macroeconomic Performance in Frankfurt, Goethe University, and at UCL, at ESWM 2020, NBER SI MDMM 2020, Konstanz Seminar 2022. We acknowledge financial support from the French National Research Agency (ANR-20-CE26-0018 IRMAC). LeGrand: Rennes School of Business and ETH Zurich; `francois.le-grand@rennes-sb.com`. Martin-Baillon: SciencesPo; `alais.martinbaillon@sciencespo.fr`. Ragot: SciencesPo, OFCE, and CNRS; `xavier.ragot@sciencespo.fr`.

1 Introduction

Monetary policy generates redistributive effects through various channels, which have been studied in a vast empirical and theoretical literature, reviewed below. However, it is not clear how these channels should change the conduct of monetary policy. It might be possible that monetary policy should take into account these effects to improve welfare, and thus participate in a function that is usually the domain of fiscal policy. Or, on the contrary, monetary policy could solely focus on monetary goals and let fiscal tools either dampen or strengthen the redistributive effects of monetary policy. To distinguish between these two claims, one must jointly solve for optimal fiscal and monetary policies in a realistic environment, where heterogeneity among agents generates a concern for redistribution.

To do so, we follow the so-called Bewley (1980) literature and assume incomplete insurance markets for idiosyncratic risks to be the main source of agent heterogeneity. This framework is known to be general enough to generate realistic income and wealth distributions. We further add nominal frictions, modeled as costly price adjustments. This environment has been named HANK following the seminal paper of Kaplan et al. (2018). Thanks to a methodological contribution explained below, we derive optimal monetary policy while considering three different assumptions for fiscal policy.

First, we assume that fiscal policy is optimally designed, and we jointly solve for optimal monetary and fiscal policies with commitment when five fiscal instruments are available: two linear taxes on real and nominal capital income, a linear tax on labor income, lump-sum transfers, and a riskless one-period public debt. We prove that optimal monetary policy then solely implements price stability and that inflation has no redistributive role over the business cycle. This result can be understood using the distinction made by Kaplan et al. (2018) and Auclert (2019) between the direct and indirect channels of monetary policy. Capital taxes are sufficient instruments for the planner to reproduce any allocation that can be reached with the direct channels, because it affects the returns on capital. The labor tax is sufficient to replicate any allocation reachable with the indirect effects, because it creates a wedge between labor income and the marginal productivity of labor. Thus, linear taxes on capital and labor are sufficient to ensure that monetary policy should solely focus on price stability and that fiscal policy alone deals with redistribution.

Second, we deviate from the assumption of optimal time-varying fiscal policy to consider optimal monetary policy with exogenous fiscal rules. Following the literature, we study the optimal inflation volatility when the economy is close to its long-run equilibrium, which is called the *timeless* perspective (see Woodford, 1999, for instance). As written by McCallum and Nelson (2000), this timeless perspective is the closest notion to “optimal policy making according to a rule”. The model must now be simulated to quantify the optimal departure from price stability. Considering a standard calibration, we find that inflation volatility is small even when tax rates are constant. Considering exogenous variations in tax rates, we also show that, as expected, the volatility of the economy can be reduced further, because those tax variations provide insurance against aggregate shocks. We also generate a more volatile inflation response by considering a calibration with a steeper Phillips curve and a very unequal redistribution of firms' profits. In this case as well, simple exogenous fiscal rules reduce the volatility of the economy (inflation, aggregate consumption, and output). To understand this result and compare it with the literature, we provide simple models where optimal monetary policy can be analyzed in various environments. Indeed, our general model features additional elements, such as capital, which makes the comparison with the literature not direct. This analysis confirms both the key role of profit distribution and of the slope of the Phillips curve for generating a sizable inflation response and that simple fiscal rules can reduce the volatility of optimal inflation.

Third, to check the robustness of these results, we solve for the optimal monetary policy and optimal time-varying labor tax, for a constant exogenous capital tax (instead of the exogenous fiscal rules considered in the previous case). We show how to calibrate a specific social welfare function, for replicating the observed US fiscal system as an optimal steady-state outcome. Assuming an interior solution, we solve for optimal labor tax and inflation dynamics, and we also find that optimal inflation volatility is small.

Deriving theoretical and quantitative results for optimal policies in incomplete-market economies with aggregate shocks is a challenging task. We perform our analysis thanks to two recent methodological contributions. First, we elaborate on the Lagrangian approach of Marcet and Marimon (2019), which appears particularly well-suited for HANK economies. We introduce the notion of *net social value of liquidity* for each agent, which considerably simplifies the algebra and interpretation. Second, to simulate the model, we use a truncated representation of incomplete insurance market economies that we apply here to a monetary economy. This theory of the (uniform) truncation is first used in LeGrand and Ragot (2022a) to study optimal unemployment benefits. We construct a consistent and accurate approximation of the economy where heterogeneity depends only on a finite but arbitrarily large number of past consecutive realizations of idiosyncratic risk. We also use the refinement of this truncation method, developed in LeGrand and Ragot (2022c), which considers possiblydifferent truncation lengths and solves the curse of dimensionality of the uniform truncation. Both methods provide quantitatively similar results and we check their accuracy by comparing their outcomes to those of the standard Reiter (2009) method, which is known to be close to other methods such as Boppart et al. (2018). We also discuss empirically relevant social welfare functions, which allows us to derive the dynamics of the model around a well-defined steady state, which can be of independent interest.

Discussion of the related literature. First, our equivalence result must be related to the five transmission channels of monetary policy that have been identified in the heterogeneous-agent literature (Kaplan et al. 2018, and Auclert, 2019, among others). Monetary policy has direct effects, going through changes in real returns (Gornemann et al., 2016). The changes in returns generate a substitution effect, already present in the representative-agent new-Keynesian model (Woodford, 2003). Inflation also affects the real value of nominal assets, through a Fisher effect (Doepke and Schneider, 2006). Changes in real returns generate a wealth effect due to unhedged interest rate exposure, identified by Auclert (2019). In addition to these three direct channels, there are also two indirect effects due to the endogeneity of labor income and to the heterogeneous exposure to income variations (Coibion et al., 2017; Acharya and Dogra, 2021). If labor and capital taxes are optimally designed, then the use of these channels by monetary policy does not increase welfare. Our equivalence results have a similar flavor to that of Correia et al. (2008), who derive equivalence results considering consumption tax in representative-agent economies with no capital. Using alternative taxes, we prove similar results considering heterogeneous portfolio holdings with both real and nominal assets.

This paper also connects to the literature investigating optimal policies with incomplete markets and heterogeneous agents. A first strand of this literature relies on tractable models featuring a simple distribution of wealth, which enables identifying the trade-offs faced by optimal policies. Challe (2020) solves for optimal monetary policy in a "zero-liquidity" environment with endogenous risk. Bilbiie and Ragot (2021) study optimal monetary policy in a tractable model with limited heterogeneity and money. Bilbiie (2021) analyzes a no-trade equilibrium with two types of agents. McKay and Reis (2021) solve for optimal simple fiscal rules (automatic stabilizers) in a tractable model considering exogenous monetary policy. In a time-0 experiment, they also find a significant deviation from price stability when fiscal instruments are missing.

A second strand of the literature analyses optimal policies with more general distributions of wealth. This is especially true for the literature on optimal fiscal policy in incomplete-market and heterogeneous-agent models (Aiyagari et al., 2002; Werning, 2007; Bassetto, 2014; Açikgöz et al., 2018; or Dyrda and Pedroni, 2022, among others). In this strand, a couple of recent papers study optimal monetary policy with incomplete insurance-markets. Nuño and Thomas (2022) solve for optimal monetary policy under commitment in an economy with uninsurable idiosyncratic risk, nominal long-term bonds, and costly inflation. They propose a methodology based on the calculus of variation. They show that the optimal policy features inflation front-loading that can be sizable in a time-0 problem, but that inflation volatility is reduced in a timeless perspective. Dávila and Schaab (2022) build on this framework and solve for the optimal monetary policy in a closed economy. They also find that the optimal deviation from price stability is close to 0. Acharya et al. (2022) solve for optimal monetary policy using the tractability of the CARA-normal environment without capital. They show that a countercyclical idiosyncratic risk creates a motive for redistribution and hence for monetary policy. They focus on a time-0 problem, and deviation from price stability remains of small magnitude in their quantitative applications, except when the price-adjustment cost becomes very small. Smirnov (2022) builds on the method of Nuño and Thomas (2022) to solve for the optimal monetary policy in a setting with an occasionally binding borrowing constraint and countercyclical income risk. Importantly, Bhandari et al. (2021b) quantitatively solve for optimal monetary and fiscal policies in a new-Keynesian model with aggregate shocks. They report a significant deviation from price stability at a long horizon, which is partly immune to the time-0 bias, when fiscal instruments are missing and the initial distribution differs from the steady-state one. This result appears to be sensitive to the calibration choices regarding the slope of the Phillips curve and the distribution of firms' profits. We find however that simple fiscal rules reduce optimal inflation volatility in a similar calibration strategy. McKay and Wolf (2022) solve for the optimal monetary policy by considering a linear-quadratic policy problem. They show that household heterogeneity adds a term to the usual loss function due to distributional motives. Finally, Yang (2022) solves for the optimal monetary policy by optimizing on the coefficients of a Taylor rule in a model with three redistributive channels for inflation: an expenditure channel (households have different consumption baskets and inflation is heterogeneous across products) and the standard Fisher and earning channels. In this literature, our contribution is to characterize the dynamics of optimal monetary policy in a timeless perspective, and in a setup with capital and occasionally binding credit constraints, and with different assumptions concerning fiscal policy.

The paper is organized as follows. Section 2 presents the environment. Section 3 proves our equivalence result. Section 4 presents the tools to simulate the general model. Section 5 reports our quantitative results. Section 6 compares our results to the literature. Section 7 concludes.

2 The environment

Time is discrete, indexed by $t \geqslant 0$. The economy is populated by a continuum of agents of size 1, distributed on a segment J following a non-atomic measure ℓ: $J(\ell) = 1$. Following Green (1994), we assume that the law of large numbers holds.

2.1 Risk

The sole aggregate shock of the economy affects the technology level. We denote this risk by $(z_t)_{t\geq 0}$, and the economy-wide productivity, denoted $(Z_t)_{t\geq 0}$, is assumed to relate to z_t through: $Z_t = \exp(z_t)$. The history of aggregate risk up to period t is denoted z^t.

In addition to the aggregate shock, agents face an uninsurable idiosyncratic labor productivity shock $y \in \mathcal{Y}$. An agent i can adjust her labor supply, denoted by $l_{i,t}$, and earns the before-tax hourly wage $\tilde{w}_t$ (which depends on the aggregate shock). Therefore, her total before-tax wage amounts to $y_{i,t}\tilde{w}_t l_{i,t}$. We assume that the productivity process is a first-order Markov chain with a constant transition matrix, denoted by $(\pi_{yy'})_{y,y'}$. The share of agents with productivity y, denoted by S_y, is constant and equal to: $S_y = \sum_{\tilde{y}} \pi_{\tilde{y}y} S_{\tilde{y}}$ for all $y \in \mathcal{Y}$. Finally, a history of productivity shocks up to date t is denoted by y^t. Using transition probabilities, we can compute the measure θ_t, such that $\theta_t(y^t)$ represents the share of agents with history y^t in period t.

2.2 Preferences

In each period, the economy has two goods: a consumption good and labor. Households are expected-utility maximizers, and they rank streams of consumption $(c_t)_{t\geq 0}$ and of labor $(l_t)_{t\geq 0}$ according to a time-separable intertemporal utility function equal to $\sum_{t=0}^{\infty} \beta^t U(c_t, l_t)$, where $\beta \in (0,1)$ is a constant discount factor and $U(c,l)$ is an instantaneous utility function. As is standard in this class of models, we focus on the case where U is separable in consumption and labor and is expressed as:[1] $U(c,l) = u(c) - v(l)$, where $u : \mathbb{R}_+ \to \mathbb{R}$ and $v : \mathbb{R}_+ \to \mathbb{R}$ are twice continuously differentiable and increasing. Furthermore, u is concave, with $u'(0) = \infty$, and v is convex.

2.3 Production

The consumption good Y_t is produced by a unique profit-maximizing representative firm that combines intermediate goods $(y_{j,t}^f)_j$ from different sectors indexed by $j \in [0,1]$ using a standard Dixit-Stiglitz aggregator with an elasticity of substitution, denoted ε:

$$Y_t = \left[\int_0^1 {y_{j,t}^f}^{\frac{\varepsilon-1}{\varepsilon}} dj \right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

For any intermediate good $j \in [0,1]$, the production $y_{j,t}^f$ is realized by a monopolistic firm and sold at price $p_{j,t}$. The profit maximization for the firm producing the final good implies:

$$y_{j,t}^f = \left(\frac{p_{j,t}}{P_t} \right)^{-\varepsilon} Y_t, \text{ where } P_t = \left(\int_0^1 p_{j,t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}.$$

The quantity P_t is the price index of the consumption good. Intermediary firms are endowed with a Cobb-Douglas production technology and use labor and capital as production factors. The production technology involves that $\tilde{l}_{j,t}$ units of labor and $\tilde{k}_{j,t}$ units of capital are transformed into $Z_t \tilde{k}_{j,t}^{\alpha} \tilde{l}_{j,t}^{1-\alpha}$ units of intermediate good. At the equilibrium, this production will exactly cover the demand $y_{j,t}^f$ for the good j, which will be sold with the real price $p_{j,t}/P_t$. We denote as $\tilde{w}_t$ the real before-tax wage and $\tilde{r}_t^K$ the real before-tax net interest rate on capital – which are both identical for all firms. The capital depreciation is denoted $\delta > 0$. Since intermediate firms have market power and internalize it, the firm's objective is to minimize production costs, including capital depreciation, subject to producing the demand $y_{j,t}^f$. The cost function $C_{j,t}$ of firm j is therefore $C_{j,t} = \min_{\tilde{l}_{j,t}, \tilde{k}_{j,t}} \{(\tilde{r}_t^K + \delta)\tilde{k}_{j,t} + \tilde{w}_t \tilde{l}_{j,t}\}$, subject to $y_{j,t}^f = Z_t \tilde{k}_{j,t}^{\alpha} \tilde{l}_{j,t}^{1-\alpha}$. Denoting by $\zeta_{j,t}$ the Lagrange multiplier of the production constraint, first-order conditions imply:

$$\tilde{r}_t^K + \delta = \zeta_{j,t} \alpha \frac{y_{j,t}^f}{\tilde{k}_{j,t}} \text{ and } \tilde{w}_t = \zeta_{j,t}(1-\alpha) \frac{y_{j,t}^f}{\tilde{l}_{j,t}}. \quad (1)$$

[1] Our equivalence result does not depend on this functional form. See Section 3.2 for a general function U.

The optimum thus features a common value, denoted by ζ_t, independent of firm type j, with:

$$\zeta_t = \frac{1}{Z_t}\left(\frac{\tilde{r}_t^K + \delta}{\alpha}\right)^{\alpha}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{1-\alpha}. \tag{2}$$

The firm j's cost becomes then $C_j = \zeta_t y_{j,t}^f$, which is linear in the demand $y_{j,t}^f$. Following the literature, we assume the presence of a subsidy τ^Y on the total cost, which will compensate for steady-state distortions, such that the total cost supported by firm j is $\zeta_t y_{j,t}^f(1-\tau^Y)$. Integrating factor price equations (1) over all firms leads to:

$$K_{t-1} = \frac{1}{Z_t}\left(\frac{\tilde{r}_t^K + \delta}{\alpha}\right)^{\alpha-1}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{1-\alpha} Y_t \text{ and } L_t = \frac{1}{Z_t}\left(\frac{\tilde{r}_t^K + \delta}{\alpha}\right)^{\alpha}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{-\alpha} Y_t, \tag{3}$$

where Y_t is the total production, with which (3) verifies:

$$Y_t = Z_t K_{t-1}^{\alpha} L_t^{1-\alpha} = \frac{(\tilde{r}_t^K + \delta)K_{t-1} + \tilde{w}_t L_t}{\zeta_t}. \tag{4}$$

Finally, in this setup, the usual factor price relationships do not hold, but we still have:

$$\frac{K_{t-1}}{L_t} = \frac{\alpha}{1-\alpha}\frac{\tilde{w}_t}{\tilde{r}_t^K + \delta}. \tag{5}$$

In a real setup (featuring $\zeta_t = 1$ for all t), equations (2) and (5) fall back to the standard definitions of factor prices: $\tilde{r}_t + \delta = \alpha Z_t(\frac{K_{t-1}}{L_t})^{\alpha-1}$ and $\tilde{w}_t = (1-\alpha)Z_t(\frac{K_{t-1}}{L_t})^{\alpha}$.

In addition to the production cost, intermediate firms face a quadratic price adjustment cost *à la* Rotemberg (1982) when setting their price. Following the literature, the price adjustment cost is proportional to the magnitude of the firm's relative price change and equal to $\frac{\kappa}{2}\left(\frac{p_{j,t}}{p_{j,t-1}} - 1\right)^2 Y_t$, where $\kappa \geq 0$. We can thus deduce the real profit, denoted $\Omega_{j,t}$, at date t of firm j:

$$\Omega_{j,t} = \left(\frac{p_{j,t}}{P_t} - \left(\frac{\tilde{r}_t + \delta}{\alpha}\right)^{\alpha}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{1-\alpha}\frac{1-\tau^Y}{Z_t}\right)\left(\frac{p_{j,t}}{P_t}\right)^{-\varepsilon} Y_t - \frac{\kappa}{2}\left(\frac{p_{j,t}}{p_{j,t-1}} - 1\right)^2 Y_t - t_t^Y, \tag{6}$$

where t_t^Y is a lump-sum tax financing the subsidy τ^Y. Computing the firm j's intertemporal profit requires to define the firm's pricing kernel. In a heterogeneous agent economy, there is no obvious choice for the pricing kernel. We discuss this aspect in Section 2.7. For the moment, we assume only that the firm's j pricing kernel is independent of its type, and we denote it by $\frac{M_t}{M_0}$. With this notation, the firm j's program, consisting in choosing the price schedule $(p_{j,t})_{t\geq 0}$ maximizing the intertemporal profit at date 0, can be expressed as: $\max_{(p_{j,t})_{t\geq 0}} \mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t \frac{M_t}{M_0}\Omega_{j,t}]$. Since this program yields a solution independent of the firm type j, all firms in the symmetric equilibrium will charge the same price: $p_{j,t} = P_t$. Denoting the gross inflation rate as $\Pi_t = \frac{P_t}{P_{t-1}}$ and setting $\tau^Y = \frac{1}{\varepsilon}$ to obtain an efficient steady state, we obtain the standard equation characterizing the

Phillips curve in our environment:

$$\Pi_t(\Pi_t - 1) = \frac{\varepsilon - 1}{\kappa}(\zeta_t - 1) + \beta \mathbb{E}_t \Pi_{t+1}(\Pi_{t+1} - 1)\frac{Y_{t+1}}{Y_t}\frac{M_{t+1}}{M_t}. \tag{7}$$

The real profit is independent of the firm's type and can be expressed as follows:

$$\Omega_t = \left(1 - \zeta_t - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t. \tag{8}$$

2.4 Assets

Agents have the possibility to trade two assets. The first one is nominal public debt, whose supply size is denoted by B_t at date t. Public debt is issued by the government and is assumed to be exempt of default risk. The nominal debt pays off a nominal gross and pre-tax interest rate that is predetermined. In other words, the nominal interest rate between dates $t-1$ and t is known at $t-1$. We denote this (gross and before tax) nominal interest rate by $\tilde{R}^N_{t-1}$. The associated real before-tax (gross) interest rate for public debt is $\tilde{R}^N_{t-1}/\Pi_t$, where Π_t is the gross inflation rate. Note that due to inflation, this ex-post real rate is not predetermined anymore. We denote by $b_{i,t}$ the debt investment of agent i. We assume that agents face nominal borrowing constraints, and their nominal debt holdings must be higher than $-\bar{b} \leq 0$. In the rest of the paper, we will focus on the case where the credit limit is above the steady-state natural borrowing limit.[2]

The second asset consists of capital shares, which pay off a (net and before-tax) real interest rate $\tilde{r}^K_t$ – as introduced above. We denote by $k_{i,t}$ the share of capital held by agent i. We assume that agents cannot issue any real claim and real asset holdings must remain positive.

2.5 Government, fiscal tools and monetary policy

In each period t, the government has to finance an exogenous public good expenditure G_t, as well as lump-sum transfers $T_t \geq 0$.[3] The latter transfers can be thought of as social transfers, which can contribute to generating progressivity in the overall tax system. Heathcote et al. (2017) have shown that such transfers are needed to properly replicate the US fiscal system. The government has several tools for financing these expenditures. First, the government can rely on four different taxes. The first two taxes are distorting taxes, denoted by τ^K_t, τ^B_t, that are levied on real and nominal asset payoffs respectively. These two taxes will be called in the remainder of the paper *real and nominal asset taxes*. The real asset tax will also be called capital tax, as is standard. Second, the government also fully taxes the firms' profits, which limits the distortions implied by profit distribution. Finally, the last tax denoted by τ^L_t concerns labor income. In addition to

[2] Aiyagari (1994) discusses the relevant values of the natural borrowing limit in an economy without aggregate shocks. Shin (2006) provides a similar discussion in the presence of aggregate shocks. A standard value in the literature is a zero borrowing limit, which ensures that consumption remains positive in all states of the world.

[3] We rule out the possibility of lump-sum taxes as a standard assumption in this literature (Aiyagari et al., 2002). See Bhandari et al. (2017) for an analysis of the case, where lump-sum taxes can be the planner's instruments.

these taxes, the government can also issue a one-period public nominal bond, and the public debt outstanding amount at date t is denoted by B_t. To sum up, fiscal policy is characterized by five instruments $(\tau_t^L, \tau_t^K, \tau_t^B, T_t, B_t)_{t\geq 0}$ given an exogenous public spending $(G_t)_{t\geq 0}$.

After-tax quantities are denoted without a tilde. The real after-tax wage w_t, as well as the real after-tax interest rates r_t^K, and R_t^N (for the capital and public debt, respectively) can therefore be expressed as follows:

$$w_t = (1-\tau_t^L)\tilde{w}_t, \tag{9}$$

$$r_t^K = (1-\tau_t^K)\tilde{r}_t^K, \quad \frac{R_t^N}{\Pi_t} - 1 = (1-\tau_t^B)(\frac{\tilde{R}_{t-1}^N}{\Pi_t} - 1), \tag{10}$$

Taxes on asset-bearing assets are asset-specific and are levied on real returns. The period-t real return on the nominal interest rate $\tilde{R}_{t-1}^N$, set in period $t-1$, is affected by both the period-t inflation rate and the period-t tax τ_t^B (hence the notation R_t^N in (10)).

The government uses its financial resources, made of labor and asset taxes, and public debt issuance, to finance public goods, lump-sum transfers, and debt repayment:

$$G_t + \frac{\tilde{R}_{t-1}^N}{\Pi_t}B_{t-1} + T_t \leq \tau_t^L \tilde{w}_t L_t + \tau_t^K \tilde{r}_t^K K_{t-1} + \tau_t^B\left(\frac{\tilde{R}_{t-1}^N}{\Pi_t} - 1\right)B_{t-1} + B_t.$$

The expression of this government budget constraint can be simplified, following Chamley (1986). Using the relationships (1) and (4), as well as the definition of post-tax rates in equations (9) and (10), the governmental budget constraint becomes:

$$G_t + \frac{R_t^N}{\Pi_t}B_{t-1} + r_t^K K_{t-1} + w_t L_t + T_t = B_t + \left(1 - \frac{\kappa}{2}(\Pi_t - 1)^2\right)Y_t - \delta K_{t-1}. \tag{11}$$

Monetary policy consists in choosing the nominal interest rate $\tilde{R}_t^N$ on public debt (between t and $t+1$), affecting the gross inflation rate Π_t. The choice of optimal monetary-fiscal policy is thus the choice of the path of the instruments $(\tau_t^L, \tau_t^K, \tau_t^B, T_t, B_t, \tilde{R}_t^N, \Pi_t)_{t\geq 0}$. These instruments are not independent of each other and are intertwined through the budget constraint of the government and the Phillips curve.

2.6 Agents' program, resource constraints, and equilibrium definition

Each agent i is endowed at date 0 by an initial real and nominal wealth $k_{i,-1}$ and $b_{i,-1}$ and an initial productive status $y_{i,0}$. At future dates, her nominal savings pay off the post-tax gross interest rate $\frac{R_t^N}{\Pi_t}$ between period $t-1$ and period t, while her real savings pay off the post-tax gross rate $1 + r_t^K$. Formally, the agent's program can be expressed, for given initial endowments

$k_{i,-1}$ and $b_{i,-1}$ and given initial shocks $y_{i,0}$ and z_0, as:

$$\max_{\{c_{i,t},l_{i,t},k_{i,t},b_{i,t}\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(c_{i,t}) - v(l_{i,t})\right), \tag{12}$$

$$c_{i,t} + k_{i,t} + b_{i,t} = (1 + r_t^K)k_{i,t-1} + \frac{R_t^N}{\Pi_t} b_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t, \tag{13}$$

$$b_{i,t} \geq -\bar{b},\ k_{i,t} \geq 0,\ c_{i,t} > 0,\ l_{i,t} > 0, \tag{14}$$

where $\mathbb{E}_0$ is an expectation operator over both idiosyncratic and aggregate risks.

Let us comment equation (13) to discuss the effect of fiscal and monetary policies in relationship to the literature. First, an unexpected change in asset tax τ_t^K or τ_t^B affects period t income, proportionally to interest payments on past real or nominal savings payoffs $\tilde{r}_t^K k_{i,t-1}$, or $(\frac{\tilde{R}_{t-1}^N}{\Pi_t} - 1)b_{i,t-1}$, respectively. Past savings payoffs are the amount of unhedged interest rate exposures (UREs) at period $t-1$, using the wording of Auclert (2019). Second, an unexpected change in inflation affects the return on the nominal holdings, $\frac{R_t^N}{\Pi_t}$, due to a Fisher effect. Third, labor tax affects the post-tax wage rate, which generates heterogeneous income and labor-supply effects. Finally, a change in the lump-sum transfer uniformly affects total income without distortion.

At date 0, the agent decides her consumption $(c_{i,t})_{t\geq 0}$, her labor supply $(l_{i,t})_{t\geq 0}$, and her saving plans $(b_{i,t})_{t\geq 0}$ and $(k_{i,t})_{t\geq 0}$ that maximize her intertemporal utility (12), subject to the budget constraint (13) and the borrowing limits and positivity constraints (14). These decisions are functions of the initial state $(b_{i,-1}, k_{i,-1}, y_{i,0})$, of the history of the idiosyncratic shock y_i^t and of the history of the aggregate shocks z^t. Thus, there exist sequences of functions, defined over $([-\bar{b}; +\infty) \times [0; +\infty) \times \mathcal{Y}) \times \mathcal{Y}^t \times \mathbb{R}^t$ and denoted by $(c_t, b_t, k_t, l_t)_{t\geq 0}$, such that an agent's optimal decisions can be written as follows:[4]

$$\begin{aligned} c_{i,t} &= c_t((b_{i,-1}, k_{i,-1}, y_{i,0}), y_i^t, z^t),\ b_{i,t} = b_t((b_{i,-1}, k_{i,-1}, y_{i,0}), y_i^t, z^t), \\ k_{i,t} &= k_t((b_{i,-1}, k_{i,-1}, y_{i,0}), y_i^t, z^t),\ l_{i,t} = l_t((b_{i,-1}, k_{i,-1}, y_{i,0}), y_i^t, z^t). \end{aligned} \tag{15}$$

In what follows, we simplify the notation and keep the i-index. For instance, we write $c_{i,t}$ instead of $c_t((b_{i,-1}, k_{i,-1}, y_{i,0}), y_i^t, z^t)$.

The first-order conditions (FOCs) corresponding to the agent's program (12)–(14) are:

$$u'(c_{i,t}) = \beta \mathbb{E}_t \left[(1 + r_{t+1}^K) u'(c_{i,t+1})\right] + \nu_{i,t}^k, \tag{16}$$

$$= \beta \mathbb{E}_t \left[\frac{R_{t+1}^N}{\Pi_{t+1}} u'(c_{i,t+1})\right] + \nu_{i,t}^b, \tag{17}$$

$$v'(l_{i,t}) = w_t y_{i,t} u'(c_{i,t}), \tag{18}$$

where the discounted Lagrange multipliers of the real and nominal credit constraints of agent i

[4] The existence of such functions is proven in Miao (2006), Cheridito and Sagredo (2016), and Açikgöz (2018).

are denoted by $\nu_{i,t}^k$ and $\nu_{i,t}^b$, respectively. Each of these Lagrange multipliers is null when agent i is not credit-constrained along the instrument under consideration.

We now express the economy-wide constraints:

$$\int_i b_{i,t}\ell(di) = B_t, \quad \int_i k_{i,t}\ell(di) = K_t, \quad \int_i y_{i,t}l_{i,t}\ell(di) = L_t, \tag{19}$$

$$\int_i c_{i,t}\ell(di) + G_t + K_t = \left(1 - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t + K_{t-1} - \delta K_{t-1}, \tag{20}$$

which correspond to the clearing of the financial markets, of the labor market, and of the goods market, respectively.[5]

Equilibrium definition. Our market equilibrium definition can be stated as follows.

Definition 1 (Sequential equilibrium) *A sequential competitive equilibrium is a collection of individual functions* $(c_{i,t}, l_{i,t}, b_{i,t}, k_{i,t}, \nu_{i,t}^b, \nu_{i,t}^k)_{t\geq 0, i\in\mathcal{I}}$*, of aggregate quantities* $(K_t, L_t, Y_t, \Omega_t, \zeta_t)_{t\geq 0}$*, of price processes* $(w_t, r_t^K, R_t^N, \tilde{w}_t, \tilde{r}_t^K, \tilde{R}_t^N)_{t\geq 0}$*, of fiscal policies* $(\tau_t^L, \tau_t^K, \tau_t^B, B_t, T_t)_{t\geq 0}$*, and of monetary policies* $(\Pi_t)_{t\geq 0}$ *such that, for an initial wealth and productivity distribution* $(b_{i,-1}, k_{i,-1}, y_{i,0})_{i\in\mathcal{I}}$*, and for initial values of capital stock and public debt verifying* $K_{-1} = \int_i k_{i,-1}\ell(di)$ *and* $B_{-1} = \int_i b_{i,-1}\ell(di)$*, and for an initial value of the aggregate shock* z_0*, we have:*

1. *given prices, the functions* $(c_{i,t}, l_{i,t}, b_{i,t}, k_{i,t}, \nu_{i,t}^b, \nu_{i,t}^k)_{t\geq 0, i\in\mathcal{I}}$ *solve the agent's optimization program in equations (12)–(14);*

2. *financial, labor, and goods markets clear at all dates: for any* $t \geq 0$*, equations (19) and (20) hold;*

3. *the government budget is balanced at all dates: equation (11) holds for all* $t \geq 0$*;*

4. *factor prices* $(w_t, r_t^K, R_t^N, \tilde{w}_t, \tilde{r}_t^K, \tilde{R}_t^N)_{t\geq 0}$ *are consistent with condition (5), as well as with post-tax definitions (9) and (10);*

5. *firms' profits* Ω_t *and the quantity* ζ_t *are consistent with equations (2) and (8);*

6. *the inflation path* $(\Pi_t)_{t\geq 0}$ *is consistent with the Phillips curve: at any date* $t \geq 0$*, equation (7) holds.*

2.7 The Ramsey problem

The goal of this paper is to determine the optimal monetary-fiscal policy that generates the sequential equilibrium-maximizing aggregate welfare, using an explicit aggregate welfare criterion.

[5] It would be equivalent to use the sequential representation and to integrate over initial states (of measure Λ) and idiosyncratic histories (of measure θ). For instance, nominal savings can be written as: $\int_i b_{i,t}\ell(di) = \sum_{y_i^t\in\mathcal{Y}^t}\sum_{y_0\in\mathcal{Y}}\int_{b_{-1}\in[-\bar{b},+\infty)}\int_{k_{-1}\in[0,+\infty)} b_t((b_{-1}, k_{-1}, y_0), y_i^t, z^t)\theta_t(y_i^t)\Lambda(db_{-1}, dk_{-1}, y_0) = B_t(z^t)$.

This is a difficult question, as the monetary-fiscal policy is composed of seven instruments $(\tau_t^L, \tau_t^K, \tau_t^B, T_t, B_t, \tilde{R}_t^N, \Pi_t)_{t\geq 0}$ which affect the saving decisions and the labor supplies of all agents, the capital stock, and the price dynamics. Interestingly, monetary policy has to balance the cost of output destruction (through price adjustment costs) and nominal debt monetization (as well as a more indirect role on mark-ups).

The aggregate welfare. We consider a utilitarian welfare function, where all agents are equally weighted. Formally, the aggregate welfare criterion can be expressed as follows:

$$W_0 := \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \left(u(c_{i,t}) - v(l_{i,t}) \right) \ell(di) \right]. \tag{21}$$

Choosing the pricing kernel. In a heterogeneous-agent economy, there is no straightforward choice for the firm's pricing kernel. We follow Acharya et al. (2022) and Bhandari et al. (2021b) and assume that firm's pricing kernel is risk-neutral. More precisely, the pricing kernel M_t is constant and normalized to 1 for all t: $M_t := 1$. As noted by others (Bhandari et al., 2021b) and as we checked ourselves, the choice of the pricing kernel has minor quantitative effects.[6]

The Ramsey program. The Ramsey program can be expressed using post-tax notation as:

$$\max_{\left(\tau_t^L,\tau_t^K,\tau_t^B,B_t,T_t,\Pi_t,w_t,r_t^K,R_t^N,\tilde{w}_t,\tilde{r}_t^K,\tilde{R}_t^N,K_t,L_t,Y_t,\Omega_t,\zeta_t,(c_{i,t},l_{i,t},b_{i,t},k_{i,t},\nu_{i,t}^b,\nu_{i,t}^k)_i\right)_{t\geq 0}} W_0, \tag{22}$$

$$G_t + \frac{R_t^N}{\Pi_t} B_{t-1} + r_t^K K_{t-1} + w_t L_t + T_t = B_t + \left(1 - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t - \delta K_{t-1}, \tag{23}$$

$$\text{for all } i \in \mathcal{I}\text{: } c_{i,t} + k_{i,t} + b_{i,t} = (1 + r_t^K) k_{i,t-1} + \frac{R_t^N}{\Pi_t} b_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t, \tag{24}$$

$$b_{i,t} \geq -\bar{b}, \nu_{i,t}^b (b_{i,t} + \bar{b}) = 0,\ \ \nu_{i,t}^b \geq 0, \tag{25}$$

$$k_{i,t} \geq 0, \nu_{i,t}^k k_{i,t} = 0,\ \ \nu_{i,t}^k \geq 0, \tag{26}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t \left[(1 + r_{t+1}^K) u'(c_{i,t+1}) \right] + \nu_{i,t}^k, \tag{27}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t \left[\frac{R_{t+1}^N}{\Pi_{t+1}} u'(c_{i,t+1}) \right] + \nu_{i,t}^b, \tag{28}$$

$$v'(l_{i,t}) = w_t y_{i,t} u'(c_{i,t}), \tag{29}$$

$$\Pi_t(\Pi_t - 1) = \frac{\varepsilon - 1}{\kappa}(\zeta_t - 1) + \beta \mathbb{E}_t \left[\Pi_{t+1}(\Pi_{t+1} - 1) \frac{Y_{t+1}}{Y_t} \right], \tag{30}$$

$$B_t = \int_i b_{i,t} \ell(di), \quad K_t = \int_i k_{i,t} \ell(di), \quad L_t = \int_i y_{i,t} l_{i,t} \ell(di), \tag{31}$$

[6] We also considered a firm's pricing kernel defined based on the weighted marginal utilities of agents: $M_t := \int_i u'(c_{i,t}) \ell(di)$. It has very small impacts on our results.

and subject to several other constraints (which are not reported here for space constraints): the definition (2) of ζ_t, the definition (4) of Y_t, the definition (8) of profits Ω_t, the factor-price relationship (5), the definitions (9)–(10) of post-tax quantities, the positivity of labor and consumption choices, and initial conditions.

The constraints in the Ramsey program include: the governmental and individual budget constraints (23) and (24), the two individual credit constraints on nominal and real asset holdings (and related constraints on $\nu^k_{i,t}$ and $\nu^b_{i,t}$) (25) and (26), Euler equations for consumption and labor (27), (28), and (29), the Phillips curve (30), and market clearing conditions for financial and labor markets (31). To simplify the derivation of first-order conditions, we use some aspects of the methodology of Marcet and Marimon (2019), which is sometimes called the Lagrangian method (Golosov et al., 2016), applied to incomplete-market environments. We denote by $\beta^t\lambda^k_{i,t}$, $\beta^t\lambda^b_{i,t}$, and $\beta^t\lambda^l_{i,t}$ the Lagrange multipliers of the Euler equations (27)–(29) of agent i at date t. Similarly, we denote by $\beta^t\gamma_t$ the Lagrange multiplier of equation (30) of the Phillips curve. The Lagrange multiplier of the government budget constraint is $\beta^t\mu_t$.

The Ramsey program (22)–(31) could be simplified further by following Chamley (1986) and removing taxes and pre-tax quantities, such that the planner directly chooses post-tax quantities. Pre-tax quantities can be deduced from their definitions (5) and (8). The definitions (9)–(10) of post-tax quantities can be used to recover tax paths from pre- and post-tax quantities. However, we will also consider economies where some fiscal instruments are kept fixed. We have thus chosen to provide the expression of the full-fledged version of the program, where fixing taxes only involves adding a constraint to the program.

As a final remark, the Ramsey program can be written in a recursive form. The state space for the planner and all agents is actually very large. It is the joint distribution over past values of Lagrange multipliers, wealth, productivity levels and the current aggregate state. This inclusion of past values of Lagrange multipliers, which makes the problem difficult, stems from the commitment of the planner not to surprise agents, such that their Euler equations hold in expectation. We skip this representation as the discussion of first-order conditions in the sequential representation is more intuitive.

3 An equivalence result

This section presents our main equivalence result and analyzes fiscal and monetary policies in different institutional setups. The analysis is greatly simplified if one introduces a new concept, the *social valuation of liquidity for agent i*, denoted by $\psi_{i,t}$, and formally defined as:

$$\psi_{i,t} := \underbrace{u'(c_{i,t})}_{\text{direct effet}} - \underbrace{\Big(\lambda_{i,k,t} - (1+r_t^K)\lambda_{i,k,t-1}\Big)\, u''(c_{i,t})}_{\text{effect on real savings}} \tag{32}$$
$$- \underbrace{\left(\lambda_{i,b,t} - \frac{R_t^N}{\Pi_t}\lambda_{i,b,t-1}\right) u''(c_{i,t})}_{\text{effect on nominal savings}} + \underbrace{\lambda_{i,l,t} y_{i,t} w_t u''(c_{i,t})}_{\text{effect on labor supply}}.$$

The valuation $\psi_{i,t}$ measures the benefit – from the planner's perspective – of transferring one extra unit of consumption to agent i. This can be interpreted as the planner's version of the marginal utility of consumption for the agent.[7] As can be seen in equation (32), this valuation consists of four terms. The first is the marginal utility of consumption $u'(c_{i,t})$, which is the private valuation of liquidity for agent i. The three other terms can be understood as the internalization, by the planner, of the economy-wide externalities of this extra consumption unit. More precisely, the second and third terms in (32) take into consideration the impact of the extra consumption unit on saving incentives from periods $t-1$ to t and from periods t to $t+1$. The interpretation of both terms is similar, except that the first one is for real savings and the second one for nominal ones. An extra consumption unit makes the agent more willing to smooth out her consumption between periods t and $t+1$, and thus makes her Euler equation (either nominal or real) more "binding". This more "binding" constraint reduces the utility by the algebraic quantity $u''(c_{i,t})\lambda_{i,x,t}$, where $\lambda_{i,x,t}$ is the Lagrange multiplier of the agent's Euler equation at date t, be it nominal with $x = b$, or real with $x = k$. The extra consumption unit at t also makes the agent less willing to smooth her consumption between periods $t-1$ and t and therefore "relaxes" the constraint of date $t-1$. This is reflected in the quantity $\lambda_{i,x,t-1}$ ($x = b, k$).

Finally, the fourth term reflects the wealth effect of the labor supply. Indeed, transferring an extra consumption unit to agent i deters her labor supply incentives through her labor supply Euler equation (18). This effect has to be internalized by the planner. Note that this term is present because we have chosen a utility function that is separable in labor and consumption. It would be absent with a GHH utility function.

In addition to $\psi_{i,t}$, another key quantity is the Lagrange multiplier, μ_t, on the governmental budget constraint. The quantity μ_t represents the marginal cost in period t of transferring one extra unit of consumption to households. Therefore, the quantity $\psi_{i,t} - \mu_t$ can be interpreted as the "net" valuation of liquidity: this is from the planner's perspective, the benefit of transferring one extra unit of consumption to agent i, net of the governmental cost. We thus define:

$$\hat{\psi}_{i,t} := \psi_{i,t} - \mu_t. \tag{33}$$

[7] To simplify the notation, we keep the index i, but the sequential representation can be derived along the lines of equation (15).

The interpretation of first-order conditions is greatly clarified by expressing them using $\hat{\psi}_{i,t}$ rather than the multiplier on Euler equations, $\lambda_{i,t}$.

3.1 The flexible-price economy

Our main result below is that the planner reproduces the flexible-price allocations, if it is possible to choose capital and labor taxes. We thus first analyze the flexible-price allocation, in which the price adjustment cost is $\kappa = 0$.[8] The Phillips curve does not constrain the planner's choices and its associated Lagrange multiplier is $\gamma_t = 0$. We can therefore follow here Chamley (1986) and express all the planner's constraints in post-tax prices. The Ramsey equilibrium can thus be derived using a narrower set of variables, which simplifies the algebra. Taxes are then recovered from the allocation. The before-tax rates $\tilde{w}_t$ and $\tilde{r}_t^K$ can be deduced from equations (2) and (5) with $\zeta_t = 1$. Taxes τ_t^L and τ_t^K are obtained from the relationships between pre-tax and post-tax rates (9) and (10). Finally, profits are null and the nominal rate $\tilde{R}_{t-1}^N$ and the nominal tax τ_t^B are undetermined.[9]

The resolution of the Ramsey program in Appendix A.2 shows that a solution to the Ramsey program is characterized by: (i) a real and nominal Euler-like equation for each individual valuation $\hat{\psi}_{i,t}$ (as long as i is unconstrained), and (ii) four first-order conditions related to the planner's four instruments.

The nominal Euler-like equation, corresponding to the first-order conditions with respect to nominal individual savings for non-credit-constrained agents, can be written as follows:

$$\hat{\psi}_{i,t} = \beta \mathbb{E}_t \left[R_{t+1}^N \hat{\psi}_{i,t+1} \right]. \tag{34}$$

Constrained agents i have no nominal Euler equation, as $b_{i,t} = -\bar{b}$ and $\lambda_{i,b,t} = 0$. Equation (34) states that the net social value of liquidity should be smoothed out over time with the post-tax nominal interest rate. It can be interpreted as a Euler-like equation for the planner and generalizes the standard individual Euler equation by taking into account the externalities of saving choices. Similarly, the real Euler-like equation, corresponding to real saving choices of unconstrained agents, can be written as follows:

$$\hat{\psi}_{i,t} = \beta \mathbb{E}_t \left[(1 + r_{t+1}^K) \hat{\psi}_{i,t+1} \right] + \beta \mathbb{E}_t \left[(1 + \tilde{r}_{t+1}^K) \mu_{t+1} \right] - \mu_t, \tag{35}$$

while for real constrained agents, we have $k_{i,t} = 0$ and $\lambda_{i,k,t} = 0$. The differences between the real and nominal Euler-like equations are twofold. First, the Euler equation (35) for $\hat{\psi}_{i,t}$ now

[8]Within the literature on optimal fiscal policy in heterogeneous agent models (Werning, 2007; Bassetto, 2014; Açikgöz et al., 2018; or Dyrda and Pedroni, 2022, among others), the concept of net social value of liquidity is new, to the best of our knowledge. It considerably eases the interpretation of the planner's first-order conditions and numerical simulations.

[9]Indeed, the government pays the before-tax rate $\tilde{R}_{t-1}^N$ and receives the nominal tax τ_t^B and both quantities have the same base, which is the previous-period public debt. Hence the government only cares about the post-tax nominal rate, as households.

involves the post-tax real interest rate instead of the nominal one, as in equation (34). Equation (34) also includes a supplementary smoothing term that involves the Lagrange multiplier μ on the governmental budget constraint. This term comes from the discrepancy between the social valuation of real savings and their cost through the governmental budget constraint. This term is absent in the nominal savings equation (34), because the the cost of governmental debt and of households' nominal savings is identical and equal to R_{t+1}^N.

The third first-order equation concerns the individual labor supply for each agent i, and it can be expressed as follows:

$$v'(l_{i,t}) + \lambda_{i,l,t} v''(l_{i,t}) = w_t y_{i,t} \left(\hat{\psi}_{i,t} + \mu_t \frac{\tilde{w}_t}{w_t} \right), \tag{36}$$

which equalizes the marginal social cost of labor (left-hand side) to its marginal benefit (right-hand side). Similarly to the expression (32) of $\psi_{i,t}$, the marginal social cost of labor involves the private cost $v'(l_{i,t})$ as well as the planner's internalization of the general-equilibrium effect of modifying individual labor supply that channels through the individual labor Euler equation (hence the presence of the multiplier $\lambda_{i,l,t}$). The marginal benefits of an extra unit of labor supply come from the related increase in individual consumption (through $\hat{\psi}_{i,t}$) and from the higher output and higher labor taxes that relax the governmental budget constraint (through μ_t).

The fourth condition, which regards the post-tax wage rate w_t, is:

$$\int_i \hat{\psi}_{i,t} y_{i,t} l_{i,t} \ell(di) = -\int_i \lambda_{i,l,t} y_{i,t} u'(c_{i,t}) \ell(di). \tag{37}$$

In the absence of any effect on the labor supply, the planner would choose the wage rate so as to set the aggregate net liquidity value – weighted by the individual labor supply in efficient terms – to zero: $\int_i \hat{\psi}_{i,t} y_{i,t} l_{i,t} \ell(di) = 0$, or equivalently, to equalize the social liquidity valuation to its marginal cost: $\int_i \psi_{i,t} y_{i,t} l_{i,t} \ell(di) = \mu_t \int_i a_{i,t-1} \ell(di)$. However, the planner has also to take into account the general-equilibrium distortions implied by wage variations that channel through all individual labor supplies. These distortions are proportional to the Lagrange multiplier on the labor Euler equation $\lambda_{i,l,t}$.

The fifth first-order condition deals with the post-tax nominal interest rate R_t^N and is:

$$\int_i \hat{\psi}_{i,t} b_{i,t-1} \ell(di) = -\int_i \lambda_{i,b,t-1} u'(c_{i,t}) \ell(di). \tag{38}$$

Similarly to equation (37), in the absence of any side effect, the planner would like to set to zero the aggregate net value of liquidity among all agents – weighted by agents' nominal asset holdings. However, the planner has also to factor in the side effect of R_t^N on nominal savings incentives, through the Euler equation. This effect is proportional to the shadow cost of the nominal Euler equation. Note that the sign of this shadow cost depends on the planner's perception of the savings quantity in the economy. It is positive when the planner perceives excess nominal savings in the economy, and negative the other way around (see LeGrand and Ragot, 2022a, for a lengthier discussion). In consequence, for instance, when there is an excess quantity of nominal savings in the economy, the total net valuation of liquidity is negative.

The sixth first-order condition regarding the post-tax real interest rate r_t^K can be expressed:

$$\int_i \hat{\psi}_{i,t} k_{i,t-1} \ell(di) = -\int_i \lambda_{i,k,t-1} u'(c_{i,t}) \ell(di), \tag{39}$$

and is the exact parallel of equation (38) but for real savings instead of nominal ones.

Finally, the last condition regarding the lump-sum transfer T_t is:

$$\int_i \hat{\psi}_{i,t} \ell(di) \leq 0. \tag{40}$$

This is an equality when $T_t > 0$, and an inequality when $T_t = 0$. Since there are no distortions implied by the lump-sum transfer, it is set such that the redistributive effect is null.

3.2 The equivalence result

The monetary economy features two complementary market imperfections. The first is the monopoly power of intermediary firms, which can yield a price markup ζ_t different from one. The second is the Rotemberg inefficiency, which prevents firms from setting their prices at no cost. The two imperfections are complementary. Indeed, in the absence of Rotemberg inefficiency (i.e., $\kappa = 0$), firms' profit maximization yields $\zeta_t = 1$ and the markup inefficiency vanishes, as can be seen from the Phillips curve in equation (7). Conversely, in the absence of imperfect competition, we have $\zeta_t = 1$, and the Phillips curve implies the Rotemberg inefficiency has no role to play. The planner's objective – in a monetary setup – therefore includes minimizing the impact of these two inefficiencies.

We now show that linear taxes on real and nominal assets, as well as on labor, are sufficient tools to offset these two inefficiencies along the business cycle, even when agents are heterogeneous. To do so, we first solve for the optimal monetary and fiscal policies when the government has access to a full set of fiscal tools. This program can be written as:

$$\max_{\left(B_t, T_t, \Pi_t, w_t, r_t^K, R_t^N, \tilde{w}_t, \tilde{r}_t^K, K_t, L_t, Y_t, \zeta_t, (c_{i,t}, l_{i,t}, b_{i,t}, k_{i,t})_i\right)_{t\geq 0}} W_0, \tag{41}$$

subject to the same equations as in the Ramsey program (22)–(31). Four observations are in order. First, we have dropped the taxes τ_t^L, τ_t^K, and τ_t^B from the Ramsey program since, as in the flexible-price case, they can be substituted by post-tax quantities w_t, r_t^K, and R^N (and recovered from allocation). Second, as in the flexible-price case again, the pre-tax nominal rate $\tilde{R}_t^N$ is also dropped, since it does not play any role. Third, the before-tax rates $\tilde{w}_t$ and $\tilde{r}_t^K$ play a role only in the markup coefficient ζ_t of equation (2) and in the factor price equation (5). In other words, the before-tax rates can be recovered from ζ_t and from the allocation. Finally, the markup coefficient only appears in the Phillips curve and can thus be recovered from the inflation path. Because the planner can optimally choose the post-tax nominal rate R_t^N, inflation only appears in the planner's program through the output destruction term (i.e., $-\kappa(\Pi_t - 1)^2/2$) in the government budget constraint. Deviating from null inflation, i.e., from $\Pi_t = 1$, has the sole consequence to shrink the feasible set of the planner. The latter therefore chooses to set the gross inflation rate to 1 at all dates, so as to neutralize the Rotemberg inefficiency. This also makes the price markup inefficiency vanish. At the end, the planner's faces the same program – and hence chooses the same allocation – as in the real economy. We summarize this first result in the following proposition.

Proposition 1 (An equivalence result) *When labor and both nominal and real asset taxes are available, the government exactly reproduces the flexible-price allocation and net inflation is null at all periods.*

The proof is in Appendix B. Proposition 1 actually holds in a more general framework, where the period utility is general (and not only separable in consumption and labor). The equivalence result would also hold with different market structures. In particular, distributing the profits to a mutual fund as in Bhandari et al. (2021b) would preserve the result if a time-varying tax on the fund payoffs is available. Indeed, a noteworthy aspect of the result is that a capital tax per asset choice is needed, hence one distinct instrument for nominal and real asset holdings.

Following the analysis of Kaplan et al. (2018) and Auclert (2019), monetary policy has direct effects, due to price changes, and indirect effects through general-equilibrium feedback. Proposition 1 states that the effects achieved by monetary policy can be achieved by labor and capital taxation. Loosely speaking, on the one hand, outcomes of the direct effects can be replicated by the linear capital tax, which globally affects the return on all savings. On the other hand, general equilibrium effects, affecting the real wage, can be replicated by the linear labor tax, which creates a wedge between the marginal labor cost of the firm, determining their pricing decision, and the labor income of households, determining their labor supply decisions. Finally, the equivalence result holds from time-0 onwards. As a consequence, it is valid both in a time-0 perspective, when instruments are chosen at the initial period before convergence to a long-run equilibrium, and in a timeless perspective, when the economy is running for a long period, such that the effect of initial conditions has vanished.

Note that the result of Proposition 1 would not hold anymore if the nominal tax is removed as an independent instrument and for instance set equal to the real capital tax. In that case inflation would be used to partly substitute for this absence of a specific nominal instrument. Formally, we could not write the program using post-tax quantities as we did, and pre- and post-tax quantities would not be independent of each other.

The result of Proposition 1 is in the same vein as Correia et al. (2008) and Correia et al. (2013), who also show that one can recover price stability if the planer has access to a time-varying consumption tax in a complete market environment. The inclusion of nominal and real asset taxes (instead of a consumption tax) allows us to connect our result to the literature on optimal capital taxation in the heterogeneous-agent literature.[10]

4 Simulating the optimal monetary policy with suboptimal fiscal policies

This section introduces additional assumptions to be able to simulate the model with an exogenous fiscal system. First, following the literature (Gornemann et al., 2016; Bhandari et al., 2021b, among many others), we introduce a risk-neutral mutual fund in Section 4.1 to remove portfolio choice. We specify the fiscal rule in Section 4.2 and present the Ramsey program in Section 4.3. Finally, Section 4.4 presents the truncation theory that is used to simulate the model.

4.1 Introducing a mutual fund

The mutual fund collects all interest-bearing asset payoffs, – i.e., the nominal ones of public debt and the real ones of the capital stock – and issues claims for households to invest in. Households can now invest in only one asset, and they are taxed on the return of the fund. All the theoretical results of the previous section, and in particular the irrelevance result of Proposition 1, remain valid with this new market structure.

The (before-tax) interest rate paid by this fund to agents is denoted by $\tilde{r}_t$. The three interest rates, for public debt, capital, and the fund, are connected by two different relationships. The first reflects the non-profit condition of the fund. We denote by A_t the total asset amount in the economy, equal to the sum of public debt and capital, which verifies: $A_t = K_t + B_t$. Since the fund holds all the public debt and the capital and sell shares, its non-profit condition at date t implies:

$$\tilde{r}_t A_{t-1} = \tilde{r}_t^K K_{t-1} + \left(\frac{\tilde{R}_{t-1}^N}{\Pi_t} - 1 \right) B_{t-1}. \tag{42}$$

The second relationship is the no-arbitrage condition between public debt holdings and capital shares. This condition states that one unit of consumption invested in each of the two assets should yield the same expected return. Formally, this condition can be written as:

$$\mathbb{E}_t \left[\frac{\tilde{R}_t^N}{\Pi_{t+1}} \right] = \mathbb{E}_t \left[1 + \tilde{r}_{t+1}^K \right]. \tag{43}$$

[10] Correia et al. (2008) analyze an economy without capital and no heterogeneity in asset holdings. In our environment with capital and heterogeneous asset holdings, we need one tax for each asset. In addition, the inclusion of capital tax may be more relevant quantitatively, at least in the US (see Trabandt and Uhlig, 2011).

Because of the fund intermediation, households make no actual portfolio choice, and we will denote by a_t their holdings in fund claims. Agents face borrowing constraints, and their fund holdings must be higher than $-\bar{a} \leq 0$.

With the introduction of the mutual fund, there is only a single tax on the interest payments of fund shares. We still denote τ_t^K this unique capital tax, and the post-tax fund interest rate is:

$$r_t = (1 - \tau_t^K)\tilde{r}_t. \tag{44}$$

With this notation, households' budget constraints and credit limits are:

$$a_{i,t} + c_{i,t} = (1 + r_t)a_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t, \tag{45}$$

$$a_{i,t} \geq -\bar{a}. \tag{46}$$

Because of the absence of portfolio choice, households have a unique consumption Euler equation:

$$u'(c_{i,t}) = \beta \mathbb{E}_t \Big[(1 + r_{t+1})u'(c_{i,t+1})\Big] + \nu_{i,t}, \tag{47}$$

where $\beta^t \nu_{i,t}$ is the Lagrange multiplier on the credit constraint (46). The governmental budget constraint and the financial market clearing conditions are:

$$G_t + B_{t-1} + r_t A_{t-1} + w_t L_t + T_t = B_t + \left(1 - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t - \delta K_{t-1}, \tag{48}$$

$$\int_i a_{i,t} \ell(di) = A_t = K_t + B_t. \tag{49}$$

4.2 The fiscal rule

We consider a fiscal system that has an affine structure, composed of lump-sum transfers and linear marginal tax rates. Such a system has often been used in the literature because it enables to simply reproduce the redistributivity of the US fiscal system, as shown for instance by Heathcote et al. (2017) and Dyrda and Pedroni (2022). Importantly, the fiscal system is not set optimally, but through ad-hoc fiscal rules. We follow the standard rules of Bohn (1998), in which the primary budget depends on the deviation of the public debt from its long-run target. Formally, we assume that taxes are defined through the following rules:

$$\tau_t^L = \tau_*^L - \sigma_1^L(Z_t - Z_*) - \sigma_2^L(Z_{t-1} - Z_*), \tag{50}$$

$$\tau_t^K = \tau_*^K - \sigma_1^K(Z_t - Z_*) - \sigma_2^K(Z_{t-1} - Z_*), \tag{51}$$

$$T_t = T_* - \sigma^T(Z_t - Z_*) - \sigma^B(B_t - B_*), \tag{52}$$

where T_*, τ_*^L, τ_*^K, B_*, and Z_* are steady-state values of the corresponding variables. We consider, as an educated guess, a simple linear rule depending on the present and past value of the technology shock. More elaborated rules could be more efficient, but these rules are simple and flexible enough for our experiment.[11]

The stationarity of public debt implies $\sigma^B > 0$. Different values for σ^T, σ_1^K, σ_2^K, σ^B, σ_1^L, and σ_2^L will correspond to rules with different insurance properties. As a benchmark, we will consider constant marginal tax rates over the business cycle, $\sigma_1^L = \sigma_2^L = 0$ and $\sigma_1^K = \sigma_2^K = 0$. We will also consider time-varying tax rates, so as to investigate the impacts of a time-varying fiscal system.

4.3 The Ramsey allocation

The Ramsey planner's program can be written as:

$$\max_{\left(w_t, r_t, \tilde{w}_t, \tilde{r}_t^K, \tilde{R}_t^N, T_t, K_t, L_t, \Pi_t, (a_{i,t}, c_{i,t}, l_{i,t}, \nu_{i,t})_i\right)_{t\geq 0}} W_0, \tag{53}$$

$$\text{s.t. } \tau_t^K, \tau_t^L, B_t \text{ following the fiscal rules (50)–(52),} \tag{54}$$

and subject to: the governmental budget constraint (48), the household budget constraint (45), the household credit constraint (46), the Euler equations on consumption (47) and labor (18) – which are unchanged, the Phillips curve (7), the market clearing condition (49), the labor market clearing condition (19), the fund no-profit condition (42), the no-arbitrage condition (43), and the post-tax rate definitions (9) and (44).

Constraints (54) imply that the dynamics of tax rates are exogenous. This is the main difference with Section 3: taxes cannot be adjusted over the business cycle to implement the optimal post-tax real wage and interest rate. For this reason, there is room for inflation to be used to redistribute wealth across agents and provide insurance against aggregate shocks. Since the algebra to determine the optimal inflation rate with fixed tax rates is more involved than in Section 3, we focus here solely on the intuitions about the trade-offs for the inflation rate. We formally derive all first-order conditions in Appendix C.

Deriving the first-order conditions of the planner for the choice of Π_t allows one to identify three mechanisms:

1. Changing inflation can affect the real wage due to the Phillips curve, which can be useful as an indirect tool to transfer resources across households.

2. Unexpected inflation specifically reduces the real return on public debt, which reduces the return on the fund and decreases real interest payments by the government.

3. Inflation destroys resources due to the adjustment cost.

These effects can be identified analytically from the program of the planner. We denote: by μ_t the Lagrange multiplier on the relevant expression of the government budget constraint, by γ_t the

[11] One could be tempted to optimize on simple rules to minimize the volatility of some targets. This is however not consistent, as the steady-state value of the taxes may not be optimal. In this case, the optimized rule would capture steady-state distortions, as discussed by Sims (2009). See Section 5.5 for a strategy to consider optimal steady-state taxes using a social welfare function.

Lagrange multiplier of the New-Keynesian Phillips curve (7), by Γ_t the Lagrange multiplier on the no-profit condition of the fund (42), and by Υ_t the Lagrange multiplier on the no-arbitrage condition (43). With this notation, the first-order condition on inflation can be written as follows:

$$\overbrace{\mu_t \kappa(\Pi_t - 1)}^{\text{Total cost due to inflation}} = \overbrace{(\gamma_{t-1} - \gamma_t)(2\Pi_t - 1)}^{\substack{\text{Manipulation of real wage} \\ \text{with NK Phillips curve}}} + \underbrace{\left(\Gamma_t(1-\tau_t^K)B_{t-1} - \beta^{-1}\Upsilon_{t-1}\right)\frac{\tilde{R}_{t-1}^N}{Y_t \Pi_t^2}}_{\text{Reducing interest payment on public debt}}.$$

4.4 Simulating the model: The truncation representation

The Ramsey problem of Section 4.3 cannot be solved with simple simulation techniques. Indeed, the Ramsey equilibrium is now a joint distribution across wealth and Lagrange multipliers, which is a high-dimensional object. The steady-state value of the set of Lagrange multipliers is not easy to find, and the planner's instruments depend on the dynamics of this joint distribution. For this reason, we use the truncation method of LeGrand and Ragot (2022a) to determine the joint distribution of wealth and Lagrange multipliers.[12] The accuracy of optimal policies has been further analyzed in LeGrand and Ragot (2022b) for both the steady state and the dynamics.

4.4.1 The uniform truncation method

The intuition of the method can be summarized as follows. In heterogeneous-agent models, agents differ according to their idiosyncratic history. An agent i has a period-t history $y_i^t = \{y_{i,0}, \ldots, y_{i,t}\}$. Let $h = (\tilde{y}_{-N+1}, \ldots, \tilde{y}_{-1}, \tilde{y}_0)$ be a given history of length N. In period t, an agent i is said to have *truncated history* h if the history of this agent for the last N periods is equal to $h = \{y_{i,t-n+1}, \ldots, y_{i,t}\}$. The idea of the truncation method is to aggregate agents having the same truncated history and to express the model using these groups of agents rather than individuals. This generated the so-called *truncated model*, which features a finite state space. In the truncated model, the agents' aggregation assumes full risk-sharing within each truncated history, while the "true" Bewley model features wealth heterogeneity among the agents having the same truncated history h. This simply comes from the heterogeneity in histories prior to the aggregation period (i.e., more than N periods ago). We capture this within-truncated-history through additional parameters – denoted by "ξs" – which are truncated-history specific. This construction yields a finite state-space representation, which is exogenous to agents' choices and

[12] Optimizing on simple rules in the spirit of Krusell and Smith (1998) is also hard to implement, as the shape of inflation is highly non-linear.

thereby allows one to compute optimal policies.[13] The details of the truncated model can be found in Section D of the Appendix.

To find the steady-state values of the Lagrange multipliers, we use the same algorithm as in LeGrand and Ragot (2022a):

1. Set a truncation length N and guess values for the planner's instruments.

2. Solve the steady-state allocation of the full-fledged Bewley model with the previous instrument values, using standard techniques.

3. Consider the truncated representation of the economy for a truncation length N.

   (a) Solve for the joint distribution of wealth and Lagrange multipliers.

   (b) Analyze whether the values of instruments are above or below their optimal value.

4. Change the instruments' values accordingly (or stop if their value is close enough to the optimal value), and redo the process from Step 2.

5. Increase the truncation length N, and restart from Step 2 until increasing N has no impact on the instruments' values.

We analyze optimal monetary policy with a given fiscal policy. This greatly simplifies the process, as it is known that the optimal gross inflation rate is $\Pi = 1$. As a consequence, the previous procedure is very fast to compute the joint distribution of wealth and Lagrange multipliers.

To use our truncation method in the presence of aggregate shocks, we further assume that:

1. The parameters $(\xi_h)_h$ remain constant and equal to their steady-state values.

2. The set of credit-constrained histories is time-invariant.

We thus assume that the time-variation of within-history heterogeneity is small enough for it to have a second-order effect. In addition, we assume that the aggregate shock is small enough for the set of credit-constrained histories not to change in the dynamics. Both assumptions rely on the choice of the truncation length and can be checked numerically.

4.4.2 The refined truncation method

The previous truncation method is simple to implement, but it has the drawback of considering many histories, some of them being very unlikely to be experienced by agents. By the law of large number, these histories concern a very small number of agents. For instance, for a truncation length of $N = 5$ used below, many histories have a size smaller than 10^{-6}. The idea

[13] Considering wealth bins is not possible, as the savings function and thus the transitions across wealth bins is endogenous to the planner's policy. This would imply a fixed point which would be very hard to solve.

of LeGrand and Ragot (2022c) is to consider different truncation lengths for different histories. For the sake of clarity, we will call this method the *refined* truncation, while the former one will be called the *uniform* truncation. Histories more likely to be experienced (i.e., with a bigger size) can be "refined", which means that they can be substituted by a set of histories with higher truncation lengths. For instance, the truncated history (y_1, y_1) $(N = 2)$ can be refined into $\{(y, y_1, y_1) : y \in \mathcal{Y}\}$, where the group of agents who have been in productivity y_1 for two consecutive period is split into $n_y (= Card \mathcal{Y})$ truncated histories.

A benefit of this construction is that the number of histories is a *linear* function of the maximum truncation length, instead of an exponential function. The construction of the refinements is detailed in LeGrand and Ragot (2022c). A difficulty of the construction that the set of refined histories must form a well-defined partition of the set of idiosyncratic histories in each period. To keep a simple exposition, we solve the model with a uniform truncation method, and only use the refinement in Section 5.4 as a robustness check of our main results.

5 Quantitative assessment

5.1 The calibration and steady-state distribution

Preferences. The period is a quarter. The discount factor is $\beta = 0.99$, and the period utility function is: $\log(c) - \chi^{-1}\frac{l^{1+1/\varphi}}{1+1/\varphi}$.[14] The Frisch elasticity of labor supply is set to $\varphi = 0.5$, which is the value recommended by Chetty et al. (2011) for the intensive margin in heterogeneous-agent models. The scaling parameter is $\chi = 0.076$, to obtain an aggregate labor supply of roughly 1/3.

Technology and TFP shock. The production function is Cobb-Douglas: $Y = ZK^{\alpha}L^{1-\alpha}$. The capital share is set to $\alpha = 36\%$ and the depreciation rate to $\delta = 2.5\%$, as in Krueger et al. (2018) among others. The TFP process is a standard AR(1) process, with $Z_t = \exp(z_t)$ and $z_t = \rho_z z_{t-1} + \varepsilon_t^z$, where $\varepsilon_t^z \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_z^2)$. We use the standard values $\rho_z = 0.95$ and $\sigma_z = 0.31\%$ to obtain a deviation of the TFP shock z_t equal to 1% at a quarterly frequency (see Den Haan, 2010, for instance).

Idiosyncratic risk. We use a standard productivity process: $\log y_t = \rho_y \log y_{t-1} + \varepsilon_t^y$, with $\varepsilon_t^y \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_y^2)$. We calibrate a persistence of the productivity process $\rho_y = 0.99$ and a standard deviation of $\sigma_y = 12.3\%$. These values are consistent with empirical estimates (Krueger et al., 2018). This process generates a realistic empirical pattern for wealth. The Gini coefficient of the wealth distribution amounts to 0.73, while the model implies an average annual capital-to-

[14] Note that matching the data constrains the Intertemporal Elasticity of Substitution (IES) parameter. Considering a log utility, we can match a Gini coefficient of wealth of 0.74 close to its empirical counterpart of 0.77. Considering a CRRA function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with $\sigma = 2$ reduces the Gini coefficient to 0.57. A higher value of σ decreases the Gini coefficient, as agents save more to self-insure.

GDP ratio of 2.5. These two values are in line with their empirical counterparts. Finally, the Rouwenhorst (1995) procedure is used to discretize the productivity process into 5 idiosyncratic states with a constant transition matrix.

Fiscal Rules. Steady-state parameters of the fiscal rules are calibrated based on the computations of Trabandt and Uhlig (2011), who use the methodology of Mendoza et al. (1994) on public finance data prior to 2008. This approach consists in computing a linear tax on capital and on labor, as well as lump-sum transfers that are consistent with the governmental budget constraint. Their estimations for the US in 2007 yield a capital tax (including both personal and corporate taxes) of 36%, a labor tax of 28% and lump-sum transfers equal to 8% of the GDP. We then consider the steady-state values $(\tau_*^L, \tau_*^K, T_*/GDP) = (28\%, 36\%, 8\%)$. This steady-state fiscal system generates two untargeted outcomes. First, it implies a public debt-to-GDP ratio equal to 63.5%, which is very close to the value of 63% estimated by Trabandt and Uhlig (2011). Second, it also implies a public spending-to-GDP ratio equal to 12.1%. This value is consistent with other quantitative investigations (Bhandari et al., 2017), even though a little bit low compared to the postwar value, which has decreased to 14.1% in 2017, from 17% in the 1970s.

Considering the dynamic part of the fiscal systems, we consider two fiscal rules, of the form (50)–(52). The first and benchmark rule (called Fiscal Rule 1) is defined by the six parameters $(\sigma_1^L, \sigma_2^L, \sigma_1^K, \sigma_2^K, \sigma^T, \sigma^B) = (0, 0, 0, 0, 8.5, 4.0)$, determined by the following constraints. First, we assume that the marginal capital and labor taxes are constant over the business cycle (hence, $\sigma_1^L = \sigma_2^L = 0$ and $\sigma_1^K = \sigma_2^K = 0$). Second, following recent empirical estimates, we match an average increase of debt over GDP of 2% after a negative TFP shock of 1%, in the range estimated by Kim and Zhang (2020) for developed countries. Finally, the rule ensures that public debt is stationary.

The second fiscal rule (called Fiscal Rule 2) is defined by the six parameters $(\sigma_1^L, \sigma_2^L, \sigma_1^K, \sigma_2^K, \sigma^T, \sigma^B) = (-0.6, 0.5, 0.6, -0.5, 8.5, 4.0)$. In this rule, labor taxes fall in recessions, whereas capital tax increases. This rule is a small deviation of Rule 1, and it will implement a similar debt path, because the reduction on labor tax is compensated by an increase in capital taxes. However, compared to Fiscal Rule 1, Fiscal Rule 2 implies different insurance properties against the aggregate risk.

Monetary parameters. We follow the literature and assume that the elasticity of substitution across goods is $\varepsilon = 6$ and the price adjustment cost is $\kappa = 100$ (see Bilbiie and Ragot, 2021, for a discussion and references). Table 1 provides a summary of the model parameters.

Steady-state equilibrium distribution. We first simulate the full-fledged Bewley model (i.e., without aggregate shocks) with the steady-state optimal inflation rate $\Pi = 1$. In Table 2, we report the wealth distribution generated by the model and compare it to the empirical

Parameter	Description	Value
	Preference and technology	
β	Discount factor	0.99
σ	Curvature utility	1
α	Capital share	0.36
δ	Depreciation rate	0.025
$\bar{a}$	Credit limit	0
χ	Scaling param. labor supply	0.076
φ	Frisch elasticity labor supply	0.5
	Shock process	
ρ_z	Autocorrelation TFP	0.95
σ_z	Standard deviation TFP shock	0.31%
ρ_y	Autocorrelation idio. income	0.99
σ_y	Standard dev. idio. income	12%
	Tax system	
τ^K	Capital tax	36%
τ^L	Labor tax	28%
T	Transfer over GDP	8%
B/Y	Public debt over yearly GDP	63.5%
G/Y	Public spending over yearly GDP	12.1%
	Monetary parameters	
κ	Price adjustment cost	100
ε	Elasticity of sub.	6

Table 1: Parameter values in the baseline calibration. See text for descriptions and targets.

distribution. We compute a number of standard statistics – listed in the first column – including the quintiles, the Gini coefficient, and Top 5% property.

The empirical wealth distribution reported in the second and third columns of Table 2 is computed using two sources, the PSID for the year 2006 and the SCF for the year 2007. The fourth column reports the wealth distribution generated by our model. The Gini coefficient of the model is 0.73, whereas it is 0.77 in the data. The model reproduces well the bottom of the distribution, but not do as well for the very top. It is known that additional model features must be introduced to match the high wealth inequality in the US, such as heterogeneous discount rates, as in Krusell and Smith (1998), or entrepreneurship, as in Quadrini (1999).

Truncation. We now construct the truncated model. We use a truncation length of $N = 5$, which with our 5 states implies $5^5 = 3125$ different truncated histories. We only focus on histories

	Data		Model
Wealth statistics	PSID, 06	SCF, 07	
Q1	−0.9	−0.2	0.0
Q2	0.8	1.2	0.2
Q3	4.4	4.6	5.2
Q4	13.0	11.9	18.1
Q5	82.7	82.5	76.4
Top 5%	36.5	36.4	33.2
Gini	0.77	0.78	0.73

Table 2: Wealth distribution in the data and in the model.

with positive size (some histories are never experienced if there is some zero probability to switch from one state to another), which reduces the number of histories to 727. In LeGrand and Ragot (2022a), it has been shown in an environment without nominal frictions that a small truncation length can provide accurate results, thanks to the introduction of the ξs parameters, as explained in Section 4.4. In Section 5.4, we show that it is also the case in the current setup and use the refined truncation as a robustness check.

5.2 Optimal inflation dynamics

We first simulate the model after a negative TFP shock of one standard deviation. The Impulse Response Functions (IRFs thereafter) are reported in Figure 1. We report IRFs for key variables and for three economies, all economies being at the steady state in period 0. The first economy (labeled Economy 1) is the benchmark economy implementing Fiscal Rule 1. The second economy (Economy 2) implements Fiscal Rule 1, but considers a constant inflation $\Pi_t = 1$ for all dates, instead of the optimal inflation path. This exercise is performed following Bhandari et al. (2021b), who note that $\Pi_t = 1$ for all t is the optimal inflation path in the complete market economy with TFP shock, due to the "divine coincidence" identified by Woodford (2003). Hence, comparing Economies 1 and 2 helps to identify the contribution of market incompleteness to the optimal inflation dynamics. The third economy (Economy 3) implements Fiscal Rule 2, where capital and labor tax rates are time-varying, and where the path of inflation is optimal. Comparing Economies 1 and 3 helps to identify the dynamic contribution of fiscal rules.

First, inflation in the benchmark economy (black line) increases on impact and then decreases to converge back to its steady-state value. The nominal rate decreases on impact. The inflation volatility is low, as the maximum change in inflation rate is 0.01%. This small change in inflation implies that the differences between the allocations of Economy 1 (optimal inflation rate) and of Economy 2 (constant inflation) are of small magnitude. Regarding the fiscal dimension, one can

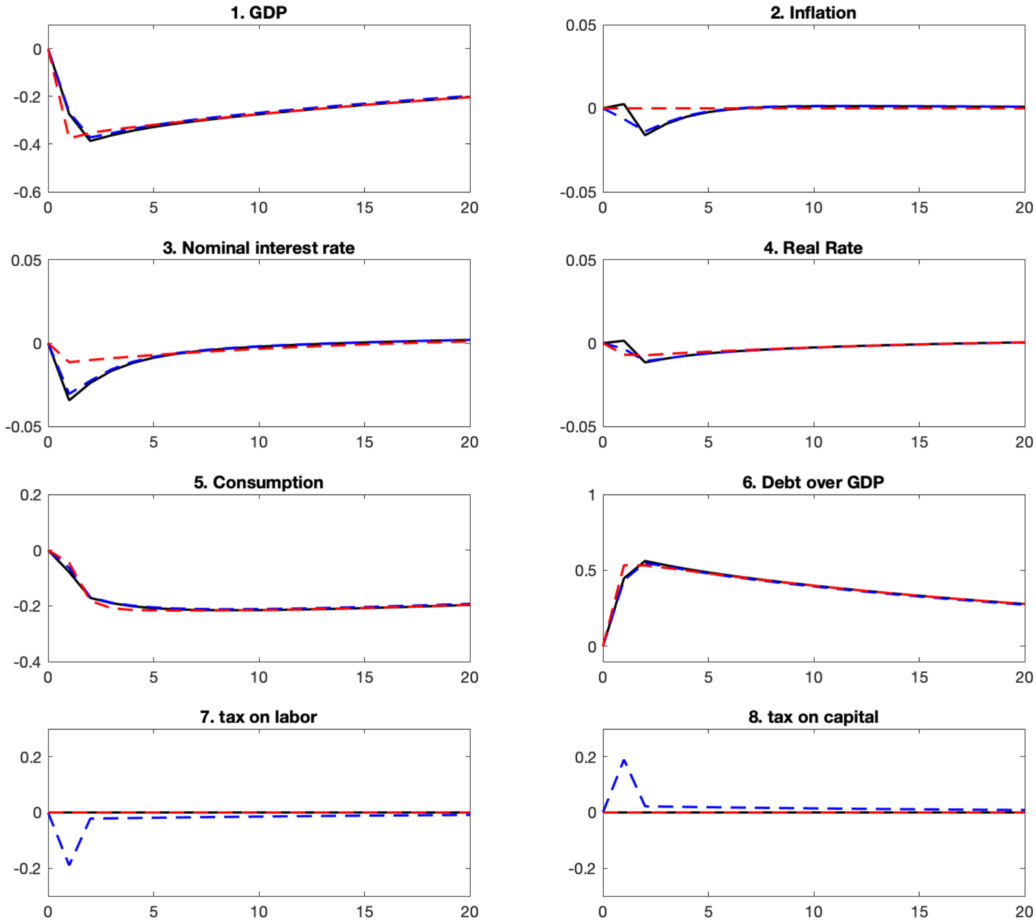


Figure 1: Impulse response functions after a negative productivity shock of one standard deviation for relevant variables. The black line is the benchmark economy with Fiscal Rule 1 and optimal monetary policy. The red dashed line is Economy 2, where we impose $\Pi_t = 1$. The blue dashed line is Economy 3 with time varying taxes (Fiscal Rule 2) and optimal monetary policy.

check that the elasticity of public debt over GDP to TFP is around -2: a fall of TFP by one standard deviation increases public debt ratio by 0.6%. Finally, tax rates are constant in this economy. The comparison of Economies 1 and 2 allows us to check that the optimal inflation path provides some insurance against the aggregate risk. This will be more clearly detailed in Section 5.3 through welfare comparisons. These results are consistent with those of Bhandari et al. (2021b), but with a lower inflation volatility due to the different calibration strategy.[15]

In Economy 3, the labor tax falls on impact by a small amount, from 28% to 27.8%, and the capital tax increases from 36% to 36.2%, while the path of public debt is roughly unchanged. Comparing Economies 1 and 3 (where tax rates are time-varying) shows that inflation (Panel 2) is slightly less volatile with Fiscal Rule 2, while aggregate consumption (Panel 5) is less volatile

[15] We verify in Section 6 below that opting for a calibration strategy similar to the one of Bhandari et al. (2021b) yields an inflation volatility that is similar to theirs.

in Economy 3 than in Economy 1, where taxes are constant. As a consequence, time-varying fiscal policy provides an efficient insurance for aggregate consumption against aggregate shocks and also contributes to reduce the volatility of the inflation response.

These findings are confirmed by second-order moments that can be found in Table 3. We simulate the economy considering the TFP process given in Table 1. We report the unconditional first- and second-order moments for the main variables, in the three economies.

		Economy 1	Economy 2	Economy 3
Y	Mean	1.43	1.43	1.43
	Std(%)	1.48	1.48	1.46
C	Mean	0.9	0.9	0.9
	Std(%)	1.33	1.34	1.31
K	Mean	14.31	14.31	14.31
	Std(%)	1.55	1.55	1.57
L	Mean	0.39	0.39	0.39
	Std(%)	0.23	0.18	0.20
τ^L	Mean	0.28	0.28	0.28
	Std(%)	0.0	0.0	0.02
τ^K	Mean	0.36	0.36	0.36
	Std(%)	0.0	0.0	0.2
B	Mean	3.64	3.64	3.64
	Std(%)	0.62	0.62	0.62
T	Mean	0.11	0.11	0.11
	Std(%)	4.74	6.0	5.26
Π	Mean	1.0	1.0	1.0
	Std(%)	0.020	0.0	0.018
Correlations				
$corr(\Pi, Y)$		0.20	0.0	0.19
$corr(\tau^K, Y)$		0.0	0.0	-0.96
$corr(\tau^L, Y)$		0.0	0.0	0.96
$corr(B, Y)$		-0.97	-0.97	-0.97
$corr(C, Y)$		0.95	0.94	0.92
$corr(Y, Y_{-1})$		0.98	0.97	0.98
$corr(B, B_{-1})$		0.96	0.97	0.96

Table 3: First- and second-order moments for key variables, in the three economies (Economy 1 is the benchmark; Economy 2 is the economy with constant inflation; Economy 3 is the economy with time-varying fiscal policy).

For each variable, we report the steady-state value (labeled "Mean") and the normalized standard deviation in percent, equal to the standard deviation divided by the mean (labeled "Std"), except for taxes and inflation, for which the standard deviation is reported. The second part of the table reports correlations. We can observe that Table 3 confirms the IRFs regarding aggregate variables. The volatility of aggregate consumption is slightly lower in Economy 1 than in Economy 2. Finally, the volatilities of inflation, aggregate consumption, and output are slightly smaller with a time-varying fiscal rule (Economy 3) than with constant taxes (Economy 1).

5.3 Welfare comparison

The welfare difference between Economies 1 and 2 is analyzed using consumption equivalents, computed in the following way. We consider two economies that differ according to their fiscal and monetary policies. The benchmark policy is denoted B, while the alternative one is denoted A. The average welfare gain is defined as the constant percentage increase in consumption of the benchmark economy, which equalizes the intertemporal welfare of the two economies. Formally, using the utilitarian social welfare function (21), the welfare gain Δ is defined as:

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\int_i\left(u((1+\Delta)c_{i,t}^B)-v(l_{i,t}^B)\right)\ell(di)\right]=\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\int_i\left(u(c_{i,t}^A)-v(l_{i,t}^A)\right)\ell(di)\right], \quad (55)$$

where the superscript A (resp. B) denotes the allocation under policy A (resp. B). If the alternative policy is welfare-improving, then $\Delta > 0$.

In our environment, and following the discussions of Bhandari et al. (2021a) and Dyrda and Pedroni (2022), this computation of welfare gains captures insurance effects against aggregate risk. Indeed, we compute welfare gains simulating the effect of two policies in the same economy, having the same steady state and hit by the same shocks. The fiscal policies do not change the steady state, and do not implement permanent transfers across agents. As a consequence, the welfare gains come from the ability of the different policies to provide insurance against aggregate risks.[16]

To better analyze the insurance property, we can condition the welfare gains on the productivity of each agent. We thus compute the welfare gain Δ_y ($y \in \mathcal{Y}$) conditional on an initial productivity level y as follows:

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\int_i(u((1+\Delta_y)c_{i,t}^B)-v(l_{i,t}^B))1_{y_{i,0}=y}\ell(di)\right]=\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\int_i(u(c_{i,t}^A)-v(l_{i,t}^A))1_{y_{i,0}=y}\ell(di)\right], \quad (56)$$

where $1_{y_{i,0}=y} = 1$ if the initial state $y_{i,0}$ is y and 0 otherwise. Consistently with the timeless perspective, we implement this welfare comparison assuming that the economy starts from the steady-state distribution. We then simulate the economy for 10,000 periods with different

[16]Following Floden (2001) and Benabou (2002), the literature has often decomposed the total welfare into three effects: a level effect, an insurance effect, and distributive effects. As we do not consider transitions starting from an initial distribution (as Bhandari et al., 2021a, or Dyrda and Pedroni, 2022), we only consider the insurance effects.

realizations of the aggregate shock history. The intertemporal welfare gains are computed as an average over these realizations, where we track agents with a given initial productivity level (which is then evolving according to the stochastic process for the idiosyncratic risk). As discussed above, the difference in the welfare gains according to the initial productivity level should be understood as a decomposition of the aggregate gains of insurance across initial types.

We find that implementing the optimal monetary policy – measured as the welfare gap between Economies 1 and 2 – generates an average welfare gain of 0.002%. This small value can be expected from Figure 1 and from the second-order moments of Table 3. However, this value hides heterogeneity among productivity levels, as can be seen in Figure 2, which plots the welfare gains between Economies 1 and 2 for different productivity levels. First, all initial productivity levels are experiencing a modest but positive welfare gain due the optimal monetary policy. The low productivity agents are experiencing a higher welfare gain, equal to 0.01%, as they enjoy a higher reduction in consumption volatility. At the opposite end, the welfare gain of very productive agents is almost null, reflecting that these agents are able to self-insure.

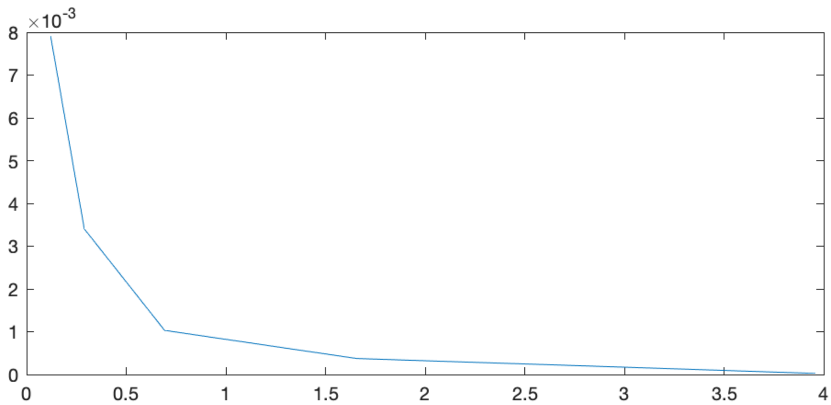


Figure 2: Welfare comparison by initial productivity, reported as the percentage increase in consumption.

5.4 Numerical robustness

The previous simulations were based on a truncation length of $N = 5$. As explained in Section 4.4.2, we now use LeGrand and Ragot (2022c) to consider heterogeneous truncation lengths. We set the truncation length $N_{\max} = 20$ for the histories with the largest size. The results are provided in Appendix E.2. They are very similar to the results with the simple truncation structure. In particular, the volatility of aggregate variables are almost unchanged, and the volatility of inflation remains very low.

Finally, we also compare the simulation outcomes of the truncation method to those of another standard method, the histogram method developed by Rios-Rull (2001), Reiter (2009), and Young (2010). We simulate the model with the previous calibration, assuming no public spending with the two methods. We then compare their IRFs, their second-order moments, as well as average and maximum absolute differences. The results are reported in Appendix F and appear to be very close.

5.5 Considering constrained-optimal fiscal policy

Finally, and as a robustness check, we also simulate the model for a constrained-optimal fiscal policy, assuming that only some fiscal tools are fixed in the dynamics, whereas other tools are optimally time-varying. To do so, we first assume that the steady-state fiscal policy is optimal for the planner, and then we compute the optimal dynamics around this optimal steady state. To obtain an optimal steady state, we assume that the planner considers an aggregate welfare criterion with Pareto weights that are history-specific. More specifically, we consider a general welfare function that depends on the weights on the utility of each agent. For the sake of generality, we assume that these weights are consistent with the sequential representation and depend on initial conditions and an idiosyncratic history. The weight of agent $i \in \mathcal{I}$ at date t is $\omega_{i,t} := \omega_t(y_i^t)$, and the weights satisfy $1 = \int_i \omega_{i,t}\ell(di) = \sum_{y_i^t \in \mathcal{Y}^t} \omega_t(y_i^t)\theta_t(y_i^t)$ for $t \geq 0$. These history-dependent weights are now also used by McKay and Wolf (2022) and Dávila and Schaab (2022). Formally, the aggregate welfare criterion can be expressed as follows:

$$W_0 = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \omega_t(y_i^t) U(c_{i,t}, l_{i,t}) \ell(di) \right]. \tag{57}$$

Following the inverse optimal taxation approach, we estimate the Pareto weights such that the observed US fiscal system is optimal at the steady state. We then solve for the optimal inflation and fiscal policy holding the capital tax constant or the labor tax constant. We find that the optimal inflation rate is almost constant, with a standard deviation of 0.01%, confirming the result that the volatility of inflation is low when fiscal policy is partly time-varying. We derive all the algebra and solution techniques in Section H.

6 Relationship with the literature: Theoretical and quantitative investigations

Despite the possible theoretical room for time-varying inflation to increase welfare in our quantitative analysis of Section 5, our baseline calibration implies a low inflation that is further reduced by a time-varying fiscal policy. As our analysis adds new ingredients to the existing literature (which are binding credit constraints, capital accumulation, a fund, fiscal policy in the timeless perspective, and the resolution via the truncation method), we now investigate how our results compare to those of Acharya et al. (2022), Bhandari et al. (2021b), or Nuño and

Thomas (2022), among others. A first difference is that we consider a timeless perspective and not a time-0 program. As explained in Nuño and Thomas (2022), inflation deviation after a TFP shock is larger in a period-0 problem compared to a timeless perspective, as the planner can surprise agents with an increase in inflation in period 0, to redistribute wealth across agents.[17]

However, some papers, such as Bhandari et al. (2021b), find that inflation can increase after a TFP shock in an environment with exogenous fiscal policy, which is immune to the time-0 inconsistency. To understand these results and the role of time-varying fiscal policy, we proceed in two steps. First, we consider in Section 6.1 a simple model, with only two types of agents. It allows us to identify the key determinants driving the inflation response, such as an unequal profit distributions, a high slope of the Phillips curve, a low IES of agents, and the absence of exogenous time-varying fiscal policy. An additional gain of the simple model is that it can be easily related to the quantitative model, with a small truncation length ($N = 1$).

Second, we verify in Section 6.2 that the factors identified in the simple model remain valid in a quantitative model (similar to the one of Section 5). This section therefore reconciles our results with those of the literature and confirms that the small inflation volatility of Section 5 is not driven by the introduction of the fund nor by the truncation method.

6.1 A two-agent economy

We consider a simplified version of the quantitative model of Section 4, with a modification of the way profits are distributed. We remove capital and consider two productivity levels $y_1 < y_2$. Agents are also assumed to inelastically supply one unit of labor and to trade nominal assets only. Productivity levels are assumed to be chosen for the aggregate labor supply to be equal to 1: $\frac{y_1+y_2}{2} = 1$ and the transition matrix (π_{ij}) is symmetric. There is no government (hence, no public spending, no tax, and no public debt). Nominal assets are thus in zero-net supply.

The firms' profits, Ω_t, are not taxed, but are distributed to households as a function of their productivity. More precisely, agents of type i receive a share of profits Ω_t equal to $(\frac{y_1^\nu+y_2^\nu}{2})^{-1}y_i^\nu$, where $\nu \geq 0$ characterizes how unequal the profit distribution is (note that shares sum to 1). When $\nu = 0$, all agents receive the same share of total profits, and the higher ν, the more the profit distribution is tilted toward high-productivity agents. At the limit ($\nu \to \infty$), all profits are distributed to the highest-productivity agents.[18]

Furthermore, this simple economy is characterized by a specific risk-sharing arrangement, which corresponds to the "island" metaphor.[19] Agents with the same productivity level pool

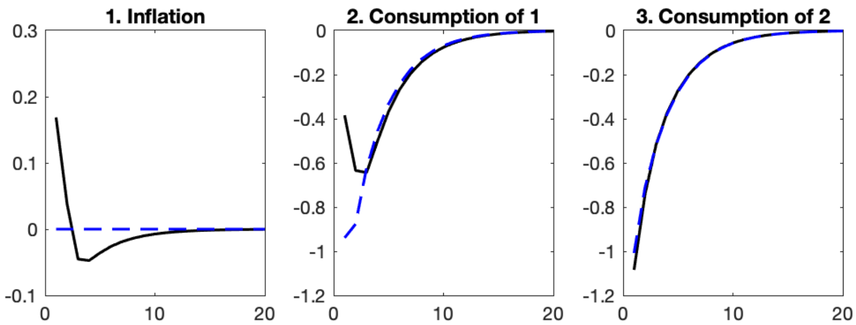
[17] In a previous version of the current paper, we quantified this time-inconsistency and related it to the period-0 capital tax problem.

[18] This flexible form allows one to encompass the case where profits are equally distributed to households (Acharya et al., 2022) and the case of an exogenous (unequal) distribution (Bhandari et al., 2021b) without introducing another asset.

[19] The simple island metaphor is used by Lucas (1990), the family metaphor is also used in Challe et al. (2017) and Bilbiie and Ragot (2021) or Bilbiie (2021). See Ragot (2018) for an overview of limited-heterogeneity models.

resources and consume the same amount, as if they were on the same island. Agents move across the two islands (corresponding to the two productivity levels) according to their productivity status. Formally, this economy is similar to a truncated model in which the truncation length is set equal to one and the ξs are all set to one. The equilibrium has indeed a simple structure with two agents (with high and low productivity), which allows us to easily derive the optimal policy with standard methods. The calculations and model equations can be found in Appendix I.1. We here simply provide simulation outcomes to identify key parameters.

A. Economy without capital



B. Economy with capital and fiscal system

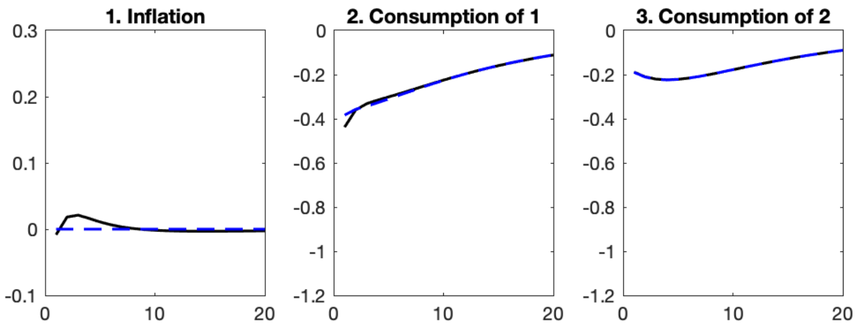


Figure 3: Optimal inflation rate (in %) and consumption levels of the two types of agents in percentage deviation, for two economies. The black solid line is the optimal allocation. The blue dashed line is the allocation when $\Pi = 1$, which is the optimal inflation in the corresponding Representative Agent (RA) economy.

The model calibration is close to the one of Bhandari et al. (2021b). The period is a year. The preference parameters are set to $\beta = 0.96$ for the discount factor and to $\sigma = 2$ for the inverse of the intertemporal elasticity of substitution. The parameters characterizing the Phillips curve are set to $\varepsilon = 6$ (to have a markup of 20%) and to $\kappa = 20$, which generates a slope of the Phillips curve equal to 6%. The productivity process is based on an AR(1) process with a persistence of $\rho_\theta = 0.992$ and a standard deviation of the innovation of $\sigma_\theta = 10.3\%$. Using the Rouwenhorst (1995) procedure, we obtain the two productivity levels $y_1 = 0.11$ and $y_2 = 0.89$ and a transition matrix defined by $\pi_{11} = \pi_{22} = 0.996$ and $\pi_{12} = \pi_{21} = 1 - \pi_{11}$. We set the parameter driving the inequality of profits distribution to $\nu = 10$. This high value implies that the high-productivity agents get almost all the profits.[20] Finally, we set the credit constraint to $\bar{b} = 0.05$, which is half of the yearly income of low-productivity agents. This value implies that low-productivity agents are credit constrained in equilibrium. We solve for the optimal monetary policy after a negative TFP shock of 1%,with a persistence of $\rho_z = 0.73$, as in Bhandari et al. (2021b).

As summary statistics, we report in Panel A of Figure 3 the inflation rate in percent and the consumption of agents 1 and 2, both in percentage deviation from their steady-state value. To understand the optimal allocation, we report the optimal response in black solid lines, and in blue dashed lines, the response with a constant inflation rate $\Pi = 1$, which corresponds to the optimal inflation rate in a Representative-Agent (RA) economy. The inflation increases at impact by around 0.17%, which is consistent with Bhandari et al. (2021b). The increase in inflation benefits to low-productivity agents (agents 1) due to the Fisher effect and to the increase in the real wage. Agents 2 experience a slightly higher fall in consumption due to inflation. Overall, inflation acts as risk-sharing device for the aggregate TFP shock.

Parameter value	Description	Inflation on impact (in%)
Benchmark		0.17
$\kappa = 100$	Price adj. cost	0.11
$\sigma = 1$	IES	0.07
$\bar{b} = 0.1$	Credit limit	0.14
$\nu = 0$	lump-sum profit red.	-0.13

Table 4: Change in inflation on impact for different parameter values.

We conduct sensitivity analysis and report in Table 4 the increase in inflation at the time of impact for different parameter choices (the shape of the inflation response being mostly unchanged). The variation of the first inflation driver cost (κ increasing from 20 to 100) implies a decrease in the slope of the Phillips curve from 6% to the standard value of 2%, which is chosen in the calibration of Section 5. The increase in inflation on impact diminishes from 0.17% to 0.1%, consistent with a higher cost for inflation. With a higher IES (equal to $1/\sigma$), the welfare benefit of consumption smoothing is reduced, which implies a lower increase in inflation, by 0.07% instead of 0.17%. A higher credit limit $\bar{b}$ increases the nominal outstanding debt of agents 1, which serve as a "fiscal" base of inflation. Finally, when the profits are equally distributed across agents (which corresponds to $\nu = 0$), inflation decreases on impact by -0.13% instead of increasing by 0.17%. Indeed, decreasing inflation allows the planner to increase firms' profits. Since the profits are equally distributed, this acts as a progressive transfer to type-1 agents (who

[20]The share of profits held by the type-1 agents is around $8 \cdot 10^{-10}$. A similar assumption is made in a number of papers (Challe et al., 2017, or Tobias et al., 2020, and see the discussion in Bilbiie, 2021).

are credit-constrained), which boosts their consumption. The sensitivity in a heterogeneous-agent economy of optimal inflation to profits distribution has already been documented (Acharya et al., 2022; Tobias et al., 2020, Bhandari et al., 2021b; Bilbiie and Ragot, 2021 or Bilbiie, 2021).

Capital and taxes in the simple model. We now introduce capital and a time-varying capital tax. The production function becomes $Y_t = Z_t K_{t-1}^{\alpha}$ with $\alpha = 1/3$ (as $L_t = 1$), and the annual depreciation rate is set to $\delta = 0.1$. Production factors are determined as in Section 2.3 together with $L_t = 1$. In addition to nominal assets, agents can also save in capital claims paying off the net interest rate $\tilde{r}_t^K$. The two-agent economy and the binding credit-constraints allow us to directly pin down the portfolio allocation, without introducing a fund. Type-2 agents hold the whole capital and the whole nominal debt issued by type-1 agents. Type-1 agents issue as much nominal debt as they can given their credit constraint. The tax system comprises a lump-sum transfer, denoted by T_t, financed by a distorting tax τ_t^K, levied on both nominal and real asset-holding payoffs. We assume that the capital tax is countercyclical and follows the simple rule $\tau_t^K = -\sigma^k z_t$, with $\sigma^k = 0.5$. This rule implies that the capital tax, as well as the lump-sum transfers, are null at the steady state, positive in recession, but negative in booms. We solve for the optimal monetary policy in this environment. To save some space, the equations are provided in Appendix I.2.

The outcome is reported in Panel B of Figure 3. Inflation barely moves on impact and in the dynamics. Time-varying redistribution across agents is now done by the tax system, and the inflation now has a modest role in the distribution of resources.

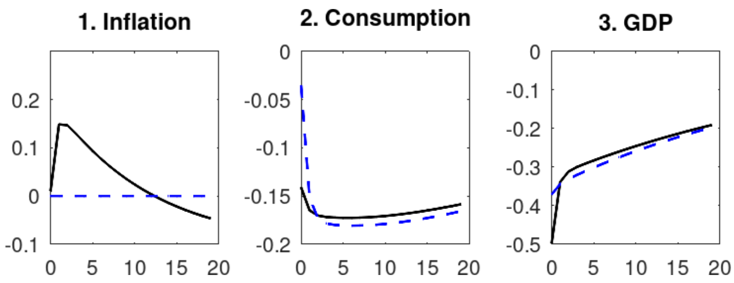
We can draw two main conclusions from our simple model exercise. The first is that monetary policy can play an active role as a risk-sharing tool (see Panel A of Figure 3). The optimal inflation response is nevertheless sensitive to the calibration, and especially to the inequality of profit distribution, the slope of the Phillips curve, and the IES. Second, introducing a simple taxation scheme that allows for redistribution basically turns off monetary policy. The inflation response becomes much smaller in the presence of fiscal tools (see Panel B of Figure 3). In other words, inflation allows for redistribution, but is a very costly substitute for fiscal policy.

6.2 Quantitative validation with an alternative calibration

We verify that the conclusions of Section 6.1, remain valid in a more quantitative model that is similar to the one of Sections 4 and 5, with the notable exception that profits are directly distributed to agents and not taxed away by the government. As in Section 6.1, agents of type i receive a share of profits Ω_t equal to $(\sum S_y y)^{-1} y_i^{\nu}$, where $\nu \geq 0$ drives how unequal the profit distribution is. Compared to Section 4, this modifies the budget constraints of the government and of households, and hence the derivation of the Ramsey program. The equations and their derivations can be found in Appendix G.

We modify the calibration of Section 5 for the factors that have been identified to be key for inflation response (see Table 4). First, the slope of the Phillips curve is steeper (as $\kappa = 20$ instead of $\kappa = 100$). Second, the IES of agents is set to 1.4 instead of 1.0. Finally, the profits distribution is very unequal ($\nu = 10$), such that they are mostly given to the most productive agents. The rest of the calibration (including for idiosyncratic and aggregate risks) is unchanged.[21]

A. Economy without fiscal rule



B. Economy with a time-varying fiscal rule

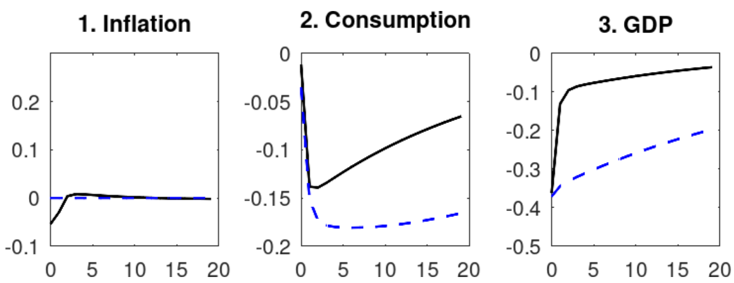


Figure 4: Optimal inflation rate (in %), aggregate consumption and GDP in percentage deviation, for two economies. The black solid line is the optimal allocation. The blue dashed line is the allocation when $\Pi = 1$, which is the optimal inflation in the corresponding RA economy.

We report the model simulation in Figure 4, where we also report in blue dashed line the outcome when $\Pi = 1$. Panel A makes it clear that the alternative calibration implies a higher response of inflation, since it approximately increases by 0.15% on impact. The response is close to the one observed in the simple model of Section 6 with no capital and no fiscal rule and also in line with Bhandari et al. (2021b). We also introduce a countercyclical fiscal policy, with the same rules as in equations (50)–(52). The results are reported in Panel B of Figure 4. Similarly to Sections 5 and 6.1, the fiscal rule reduces inflation volatility, here by one order of magnitude, from 0.68% to 0.07%. Inflation response on impact is also reduced from 0.15% to less than 0.01%. Furthermore, aggregate consumption is also less volatile, showing that the fiscal rule provides more insurance against the aggregate risk. The detailed model outcomes, including second-order

[21] To the best of our knowledge, the calibration of Section 5, in particular the slope of the Phillips curve, is more standard.

moments, can also be found in Appendix G.

7 Conclusion

We derive optimal monetary policy with commitment in an economy with incomplete insurance markets, nominal frictions, and aggregate shocks in a timeless perspective. We consider three different fiscal regimes. First, when linear capital and labor taxes can optimally vary over the business cycle, we find that there is no role for monetary policy to deviate from price stability. Redistribution is only a matter of fiscal policy. Second, when taxes are set by an exogenous fiscal rule, the inflation response is modest when the calibration is standard and further reduced when capital and labor taxes are (exogenously) time-varying. A sizable inflation response, in the absence of time-varying taxes, can be obtained when considering a combination of unequal profit distribution and a calibration with a steep Phillips curve and a low IES. Third, we consider constrained-optimal fiscal policies, and we find that there is no significant quantitative deviation from price stability. As a conclusion, inflation appears to have a modest role to play for redistribution and can efficiently be replaced by simple time-varying fiscal rules.

References

Açikgöz, O., M. Hagedorn, H. Holter, and Y. Wang (2018): "The Optimum Quantity of Capital and Debt," Working Paper, University of Oslo.

Açikgöz, O. T. (2018): "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production," *Journal of Economic Theory*, 173, 18–55.

Acharya, S., E. Challe, and K. Dogra (2022): "Optimal Monetary Policy According to HANK," Working Paper, Federal Reserve Bank of New York.

Acharya, S. and K. Dogra (2021): "Understanding HANK: Insights from a PRANK," *Econometrica*, 88, 1113–1158.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659–684.

Aiyagari, S. R., A. Marcet, T. J. Sargent, and J. Seppälä (2002): "Optimal Taxation without State-Contingent Debt," *Journal of Political Economy*, 110, 1220–1254.

Auclert, A. (2019): "Monetary Policy and the Redistribution Channel," *American Economic Review*, 109, 2333–2367.

Bassetto, M. (2014): "Optimal Fiscal Policy with Heterogeneous Agents," *Quantitative Economics*, 5, 675–704.

Benabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70, 481–517.

Bewley, T. F. (1980): "The Optimum Quantity of Money," in *Models of Monetary Economies*, ed. by J. Kareken and N. Wallace, Federal Reserve Bank of Minneapolis, 169–210.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2017): "Public Debt in Economies with Heterogeneous Agents," *Journal of Monetary Economics*, 91, 39–51.

——— (2021a): "Efficiency, Insurance, and Redistribution Effects of Government Policies," Working paper, University of Minnesota.

——— (2021b): "Inequality, Business Cycles, and Monetary-Fiscal Policy," *Econometrica*, 89, 2559–2599.

Bilbiie, F. (2021): "Monetary Policy and Heterogeneity: An Analytical Framework," Working Paper, Lausanne University.

Bilbiie, F. and X. Ragot (2021): "Optimal Monetary Policy and Liquidity with Heterogeneous Households," *Review of Economic Dynamics*, 41, 71–95.

Bohn, H. (1998): "The Behavior of US Public Debt and Deficits," *Quarterly Journal of Economics*, 4, 1329–1368.

Boppart, T., P. Krusell, and K. Mitman (2018): "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative," *Journal of Economic Dynamics and Control*, 89, 68–92.

Challe, E. (2020): "Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy," *American Economic Journal: Macroeconomics*, 2, 241–283.

Challe, E., J. Matheron, X. Ragot, and J. F. Rubio-Ramirez (2017): "Precautionary Saving and Aggregate Demand," *Quantitative Economics*, 8, 435–478.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607–622.

Cheridito, P. and J. Sagredo (2016): "Existence of Sequential Competitive Equilibrium in Krusell-Smith Type Economies," Working Paper, ETH Zurich.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–475.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017): "Innocent Bystanders? Monetary Policy and Inequality," *Journal of Monetary Economics*, 88, 70–89.

Correia, I., E. Farhi, J.-P. Nicolini, and P. Teles (2013): "Unconventional Fiscal Policy at the Zero Bound," *American Economic Review*, 4, 1172–1211.

Correia, I., J.-P. Nicolini, and P. Teles (2008): "Optimal Fiscal and Monetary Policy: Equivalence Results," *Journal of Political Economy*, 1, 141–170.

Dávila, E. and A. Schaab (2022): "Welfare Assessments with Heterogeneous Individuals," Working Paper, Yale University.

Den Haan, W. J. (2010): "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents," *Journal of Economic Dynamics and Control*, 1, 79–99.

Doepke, M. and M. Schneider (2006): "Inflation and the Redistribution of Nominal Wealth," *Journal of Political Economy*, 114, 1069–1097.

Dyrda, S. and M. Pedroni (2022): "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks," *Review of Economic Studies*, forthcoming.

Floden, M. (2001): "The Effectiveness of Government Debt and Transfers as Insurance," *Journal of Monetary Economics*, 48, 81–108.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Application*, Princeton University Press, 2nd ed.

Golosov, M., A. Tsyvinski, and N. Werquin (2016): "Recursive Contracts and Endogenously Incomplete Markets," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2, 725–741.

Gornemann, N., K. Kuester, and M. Nakajima (2016): "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy," Working Paper, Federal Reserve Bank of Philadelphia.

Green, E. (1994): "Individual-Level Randomness in a Nonatomic Population," Working Paper, University of Minnesota.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132, 1693–1754.

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

Kim, Y. and J. Zhang (2020): "The Relationship between Debt and Output," Working Paper 2020-30, Federal Reserve Bank of Chicago.

Krueger, D., K. Mitman, and F. Perri (2018): "On the Distribution of the Welfare Losses of Large Recessions," in *Advances in Economics and Econometrics: Volume 2, Eleventh World Congress of the Econometric Society*, ed. by B. Honoré, A. Pakes, M. Piazzesi, and L. Samuleson, Cambridge University Press, 143–184.

Krusell, P. and A. A. J. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

LeGrand, F. and X. Ragot (2022a): "Managing Inequality over the Business Cycle: Optimal Policies with Heterogeneous Agents and Aggregate Shocks," *International Economic Review*, 63, 511–540.

——— (2022b): "Optimal Policies with Heterogeneous Agents: Truncation and Transitions," Working Paper, SciencesPo.

——— (2022c): "Refining the Truncation Method to Solve Heterogeneous-Agent Models," *Annals of Economics and Statistics*, 146, 65–92.

Lucas, R. E. J. (1990): "Liquidity and Interest Rates," *Journal of Economic Theory*, 50, 237–254.

Marcet, A. and R. Marimon (2019): "Recursive Contracts," *Econometrica*, 87, 1589–1631.

McCallum, B. and E. Nelson (2000): "Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models," NBER Working Paper 7915, National Bureau of Economic Research.

McKay, A. and R. Reis (2021): "Optimal Automatic Stabilizers," *Review of Economic Studies*, 88, 2375–2406.

McKay, A. and C. Wolf (2022): "Optimal Policy Rules in HANK," Working Paper, FRB Minneapolis.

Mendoza, E., A. Razin, and L. Tesar (1994): "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption," *Journal of Monetary Economics*, 34, 297–323.

Miao, J. (2006): "Competitive Equilibria of Economies with a Continuum of Consumers and Aggregate Shocks," *Journal of Economic Theory*, 128, 274–298.

Nuño, G. and C. Thomas (2022): "Optimal Redistributive Inflation," *Annals of Economics and Statistics*, 146, 3–64.

Quadrini, V. (1999): "The Importance of Entrepreneurship for Wealth Concentration and Mobility," *Review of Income and Wealth*, 45, 1–19.

Ragot, X. (2018): "Heterogeneous agents in the Macroeconomy: Reduced-heterogeneity representations," in *Handbook of Computational Economics*, ed. by C. Hommes and B. Lebaron, Elsevier, forthcoming, vol. 4.

Reiter, M. (2009): "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 33, 649–665.

Rios-Rull, V. (2001): "Computation of Equilibria in Heterogenous Agent Models," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and A. Scott, Oxford University Press, chap. 11.

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

Rouwenhorst, G. K. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Structural Models of Wage and Employment Dynamics*, ed. by T. Cooley, Princeton: Princeton University Press, 201–213.

Shin, Y. (2006): "Ramsey Meets Bewley: Optimal Government Financing with Incomplete Markets," Working Paper, Washington University in St. Louis.

Sims, C. (2009): "Linearized Stochastic Models," Mimeo, Princeton University.

Smirnov, D. (2022): "Optimal Monetary Policy in HANK," Working paper, Universitat Pompeu Fabra.

Tobias, B., N.-J. Hansen, K. Per, and E. Oberg (2020): "The New Keynesian Transmission Mechanism: A Heterogeneous-Agent Perspective," *The Review of Economic Studies*, 87, 77–101.

Trabandt, M. and H. Uhlig (2011): "The Laffer Curve Revisited," *Journal of Monetary Economics*, 58, 305–327.

Werning, I. (2007): "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 122, 925–967.

Woodford, M. (1999): "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?" prepared for New Challenges for Monetary Policy, a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 26-28, 1999.

——— (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

Yang, Y. (2022): "Redistributive Inflation and Optimal Monetary Policy," Working paper, Princeton University.

Young, E. (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell-Smith Algorithm and Non-Stochastic Simulations," *Journal of Economic Dynamics and Control*, 34, 36–41.

Appendix

A Computing the FOCs of the full-fledged Ramsey program with flexible prices

This section is organized in two parts. In Section A.1, we transform the Ramsey program by including the Euler equations in the planner's objective. In Section A.2, we derive the FOCs of the Ramsey program in the flexible-price economy.

A.1 Transforming the Ramsey program

We explicitly include the Euler equations (the two on consumption and the one on labor) in the planner's objective. We recall that the three Lagrange multipliers are denoted by $\beta^t \lambda_{i,k,t}$ for the real Euler equation, $\beta^t \lambda_{i,b,t}$ for the nominal Euler equation, and $\beta^t \lambda_{i,l,t}$ for the labor Euler equation. The objective of the Ramsey program (22)–(31) then becomes:

$$
\begin{aligned}
J = {} & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i (u(c_{i,t}) - v(l_{i,t}))\ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,k,t} \left(u'(c_{i,t}) - \nu_{i,k,t} - \beta \mathbb{E}_t \left[(1 + r_{t+1}^K) u'(c_{i,t+1}) \right] \right) \ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,b,t} \left(u'(c_{i,t}) - \nu_{i,b,t} - \beta \mathbb{E}_t \left[\frac{R_{t+1}^N}{\Pi_{t+1}} u'(c_{i,t+1}) \right] \right) \ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,l,t} \left(v'(l_{i,t}) - w_t y_{i,t} u'(c_{i,t}) \right) \ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \gamma_t \left(\Pi_t(\Pi_t - 1) Y_t - \frac{\varepsilon - 1}{\kappa} (\zeta_t - 1) Y_t - \beta \mathbb{E}_t \left[\Pi_{t+1}(\Pi_{t+1} - 1) Y_{t+1} \right] \right),
\end{aligned}
$$

In the previous expression, we have used the fact that when the credit constraint is binding for agent i at period t, ($\nu_{i,k,t} > 0$), then there is no Euler equation to consider ($\lambda_{i,k,t} > 0$); in other words: $\nu_{i,k,t}\lambda_{i,k,t} = \nu_{i,b,t}\lambda_{i,b,t} = 0$. With $\gamma_{-1} = 0$, we obtain after some manipulations:

$$
\begin{aligned}
J = {} & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i (u(c_{i,t}) - v(l_{i,t}))\ell(di) - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \left(\lambda_{i,k,t} - (1 + r_t^K) \lambda_{i,k,t-1} \right) u'(c_{i,t}) \ell(di) \qquad (58) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \left(\lambda_{i,b,t} - \frac{R_t^N}{\Pi_t} \lambda_{i,b,t-1} \right) u'(c_{i,t}) \ell(di) - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,l,t} \left(v'(l_{i,t}) - w_t y_{i,t} u'(c_{i,t}) \right) \ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left((\gamma_t - \gamma_{t-1}) \Pi_t(\Pi_t - 1) - \frac{\varepsilon - 1}{\kappa} \gamma_t (\zeta_t - 1) \right) Y_t,
\end{aligned}
$$

Using (58), the Ramsey program (22)–(31) can now be expressed as: $\max J$ on $(B_t, T_t, \Pi_t, w_t, r_t^K,$ $R_t^N, \tilde{w}_t, \tilde{r}_t^K, \tilde{R}_t^N, K_t, L_t, Y_t, \zeta_t, (c_{i,t}, l_{i,t}, b_{i,t}, k_{i,t}, \nu_{i,t}^b, \nu_{i,t}^k)_i)_{t \geq 0}$, subject to the same set of constraints, except the individual Euler equations for consumption (27) and the Phillips curve (30). Note that we could also write the Ramsey program using the sequential representation (as in equations (15)), but at the cost of tedious notation.

A.2 First-order conditions for the flexible-price economy

The flexible-price program implies $\kappa = 0$, and hence $\zeta_t = \Pi_t = 1$. Profits are also null. Taking advantage of the post-tax notation, the allocation is the solution of:

$$\max_{\left(B_t, T_t, w_t, r_t^K, R_t^N, (l_{i,t}, b_{i,t}, k_{i,t})_i\right)} J$$
$$c_{i,t} = -k_{i,t} - b_{i,t} + (1 + r_t^K)k_{i,t-1} + R_t^N b_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t,$$
$$G_t + R_t^N B_{t-1} + r_t^K K_{t-1} + w_t L_t + T_t = B_t + Y_t - \delta K_{t-1},$$

and where we use $K_t = \int_i k_{i,t}\ell(di),\ B_t = \int_i b_{i,t}\ell(di),\ L_t = \int_i y_{i,t} l_{i,t}\ell(di),\ Y_t = Z_t K_t^\alpha L_t^{1-\alpha}$.

The Lagrangian can be expressed as follows:

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i (u(c_{i,t}) - v(l_{i,t}))\ell(di) - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \left(\lambda_{i,k,t} - (1 + r_t^K)\lambda_{i,k,t-1}\right) u'(c_{i,t})\ell(di)$$
$$- \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \left(\lambda_{i,b,t} - R_t^N \lambda_{i,b,t-1}\right) u'(c_{i,t})\ell(di) - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,l,t} \left(v'(l_{i,t}) - w_t y_{i,t} u'(c_{i,t})\right) \ell(di)$$
$$+ \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mu_t \left(B_t + Y_t - \delta K_{t-1} - G_t - R_t^N B_{t-1} - r_t^K K_{t-1} - w_t L_t - T_t\right).$$

Derivative with respect to w_t.

$$0 = \int_i \hat{\psi}_{i,t} y_{i,t} l_{i,t} \ell(di) + \int_i \lambda_{i,l,t} y_{i,t} u'(c_{i,t})\ell(di).$$

Derivative with respect to r_t^K.

$$\int_i \lambda_{i,k,t-1} u'(c_{i,t})\ell(di) + \int_i \hat{\psi}_{i,t} k_{i,t-1}\ell(di) = 0. \tag{59}$$

Derivative with respect to R_t^N. As in the real case, we have:

$$\int_i \lambda_{i,b,t-1} u'(c_{i,t})\ell(di) + \int_i \hat{\psi}_{i,t} b_{i,t-1}\ell(di) = 0. \tag{60}$$

Derivative with respect to $k_{i,t}$. For unconstrained agents:

$$\hat{\psi}_{i,t} = \beta\mathbb{E}_t \left[(1 + r_{t+1}^K)\hat{\psi}_{i,t+1}\right] + \beta\mathbb{E}_t \left[\left(\tilde{r}_{t+1}^K - r_{t+1}^K\right)\mu_{t+1}\right] + \beta\mathbb{E}_t \left[(1 + r_{t+1}^K)\mu_{t+1}\right] - \mu_t,$$
$$= \beta\mathbb{E}_t \left[(1 + r_{t+1}^K)\hat{\psi}_{i,t+1}\right] + \beta\mathbb{E}_t \left[(1 + \tilde{r}_{t+1}^K)\mu_{t+1}\right] - \mu_t$$

while for constrained ones we have $k_{i,t} = 0$ and $\lambda_{k,i,t} = 0$.

Derivative with respect to $b_{i,t}$. For unconstrained agents:

$$\hat{\psi}_{i,t} = \beta \mathbb{E}_t \left[R^N_{t+1} \hat{\psi}_{i,t+1} \right],$$

while for constrained ones we have $b_{i,t} = 0$ and $\lambda_{b,i,t} = 0$.

Derivative with respect to $l_{i,t}$.

$$v'(l_{i,t}) + \lambda_{i,l,t} v''(l_{i,t}) = w_t y_{i,t} \left(\hat{\psi}_{i,t} + \mu_t \frac{\tilde{w}_t}{w_t} \right).$$

Derivative with respect to T_t.

$$\int_i \hat{\psi}_{i,t} \ell(di) \leq 0, \tag{61}$$

with an equality when $T_t > 0$.

B Proof of the equivalence result of Proposition 1

The first step is to simplify the expression of the Ramsey program (22)–(31). Using post-tax notation, we can remove the taxes τ^L_t, τ^K_t, and τ^B_t from the Ramsey program since they can be recovered from post-tax and pre-tax rate w_t, r^K_t, and R^N, as well as $\tilde{w}_t$, $\tilde{r}^K_t$, and $\tilde{R}^N$ using the relationships (9)–(10). We can also remove the pre-tax nominal rate $\tilde{R}^N_t$, as it does not play any role (only the post-tax rate is determined). The before-tax rates $\tilde{w}_t$ and $\tilde{r}^K_t$ can also be taken out of the Ramsey program, as they can be recovered from the allocation and the markup ζ_t through equations (2) and (5). The profit can also be removed from planner's instruments as it does not directly appear in the program and can be deduced from its definition (8) and the allocation. The second step is that we can reformulate the Ramsey program by setting $\hat{R}_t = \frac{R^N_t}{\Pi_t}$. Since the planner chooses both R^N_t and Π_t, this change of variable has no impact on allocation. The path of $(R^N_t)_t$ can be deduced from the paths of $(\Pi_t)_t$ and $(\hat{R}_t)_t$.

We can then write the Ramsey program as follows:

$$\max_{\left(B_t,T_t,\Pi_t,w_t,r_t^K,\hat{R}_t,K_t,L_t,Y_t,\zeta_t,(c_{i,t},l_{i,t},b_{i,t},k_{i,t},\nu_{i,t}^b,\nu_{i,t}^k)_i\right)_{t\geq 0}} W_0, \tag{62}$$

$$G_t + \hat{R}_t B_{t-1} + r_t^K K_{t-1} + w_t L_t + T_t = B_t + \left(1 - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t - \delta K_{t-1}, \tag{63}$$

for all $i \in \mathcal{I}$: $c_{i,t} + k_{i,t} + b_{i,t} = (1 + r_t^K)k_{i,t-1} + \hat{R}_t b_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t,$ (64)

$$b_{i,t} \geq -\bar{b}, \nu_{i,t}^b(b_{i,t} + \bar{b}) = 0,\ \nu_{i,t}^b \geq 0, \tag{65}$$

$$k_{i,t} \geq 0, \nu_{i,t}^k k_{i,t} = 0,\ \nu_{i,t}^k \geq 0, \tag{66}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t\left[(1 + r_{t+1}^K)u'(c_{i,t+1})\right] + \nu_{i,t}^k, \tag{67}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t\left[\hat{R}_{t+1} u'(c_{i,t+1})\right] + \nu_{i,t}^b, \tag{68}$$

$$v'(l_{i,t}) = w_t y_{i,t} u'(c_{i,t}), \tag{69}$$

$$B_t = \int_i b_{i,t}\ell(di),\ K_t = \int_i k_{i,t}\ell(di),\ L_t = \int_i y_{i,t} l_{i,t}\ell(di),\ Y_t = Z_t K_{t-1}^{\alpha} L_t^{1-\alpha} \tag{70}$$

$$\Pi_t(\Pi_t - 1) = \frac{\varepsilon - 1}{\kappa}(\zeta_t - 1) + \beta \mathbb{E}_t\left[\Pi_{t+1}(\Pi_{t+1} - 1)\frac{Y_{t+1}}{Y_t}\right], \tag{71}$$

without further constraints. In the program (62)–(71), the instrument $(\zeta_t)_t$ only plays a role in the Phillips curve (71). The instrument $(\zeta_t)_t$ and the Phillips curve as the constraint can be removed from the Ramsey program and substituted by the constraint $\beta \mathbb{E}_t\left[|\Pi_{t+1}(\Pi_{t+1} - 1)\frac{Y_{t+1}}{Y_t}|\right] < \infty$, which ensures that the Phillips curve can then used to deduce the path of $(\zeta_t)_t$ from the paths of $(\Pi_t)_t$ and $(Y_t)_t$.

It then follows that the Ramsey program simplifies to the set of equations (62)–(70), without $(\zeta_t)_t$ as instrument. Inflation then only appears in the government budget constraint (63) as a cost. This makes it clear that any deviation from price stability (i.e., $\Pi_t = 1$ at all dates) shrinks the feasible set of the planner and is hence avoided by the planner. This proves Proposition 1.

To see more precisely that the planner refuses any deviation from price stability, assume that the planner optimal allocation is $(B_t, T_t, \Pi_t, w_t, r_t^K, \hat{R}_t, K_t, L_t, Y_t, (c_{i,t}, l_{i,t}, b_{i,t}, k_{i,t}, \nu_{i,t}^b, \nu_{i,t}^k)_i)_t$ and that there is a date t_0 in which the planner chooses to deviate from price stability: $\Pi_{t_0} \neq 1$. Consider the allocation $(B_t, T_t', \Pi_t', w_t, r_t^K, \hat{R}_t, K_t, L_t, Y_t, \zeta_t, (c_{i,t}', l_{i,t}, b_{i,t}, k_{i,t}, \nu_{i,t}^b, \nu_{i,t}^k)_i)_t$ that is identical to the initial one, but for the transfer, inflation and individual consumption. We define for all $t \neq t_0$: $T_t' = T_t$, $\Pi_t' = \Pi_t$, and $c_{i,t}' = c_{i,t}$ and for t_0: $T_{t_0}' = T_{t_0} + \frac{\kappa}{2}(\Pi_{t_0} - 1)^2 Y_{t_0}$, $\Pi_{t_0}' = 1$, and $c_{i,t_0}' = c_{i,t} + \frac{\kappa}{2}(\Pi_{t_0} - 1)^2 Y_{t_0} > c_{i,t}$. In words, the prime allocation is the same as the non-prime one, except at date t_0, where gross inflation has been modified from $\Pi_{t_0} \neq 1$ to 1 and the output destruction $\frac{\kappa}{2}(\Pi_{t_0} - 1)^2 Y_{t_0}$ that has been avoided has been transferred to agents via the lump-sum instrument. It is easy to check that the prime allocation is feasible (i.e., that it verifies all constraints (63)–(70)) and that it implies a strictly better aggregate welfare (labor supply is unchanged, consumption at date t_0 is strictly higher and consumption at other dates unchanged). This contradicts that the non-prime allocation is optimal and shows that the planner never chooses to deviate from price stability.

C Program with suboptimal fiscal policy

Because pre- and post-tax rates cannot be set independently, nominal frictions cannot be suppressed in this setup. We have two additional Lagrange multipliers: (i) $\beta^t \Upsilon_t$ on the no-arbitrage condition (43), and (ii) $\beta^t \Gamma_t$ on the zero-profit condition (42) of the fund. In this context, the planner's objective can be written as:

$$
\begin{aligned}
\mathcal{L} = & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i (u(c_{i,t}) - v(l_{i,t}))\ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \left(\lambda_{i,c,t} - (1+r_t)\lambda_{i,c,t-1}\right) u'(c_{i,t})\ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{i,l,t} \left(v'(l_{i,t}) - w_t y_{i,t} u'(c_{i,t})\right) \ell(di) \\
& - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\gamma_t - \gamma_{t-1}) \Pi_t (\Pi_t - 1) Y_t + \frac{\varepsilon - 1}{\kappa} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \gamma_t (\zeta_t - 1) Y_t \\
& + \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mu_t \left(B_t + Y_t (1 - \frac{\kappa}{2}(\Pi_t - 1)^2) - \delta(A_{t-1} - B_{t-1}) - G_t - B_{t-1} - r_t A_{t-1} - w_t L_t - T_t \right) \\
& + \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^{t-1} \Upsilon_{t-1} (\frac{\tilde{R}^{B,N}_{t-1}}{\Pi_t} - (1 + \tilde{r}^K_t)) \\
& + \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Gamma_t (r_t A_{t-1} - (1 - \tau^K_t)(\tilde{r}^K_t K_{t-1} + (\frac{\tilde{R}^{B,N}_{t-1}}{\Pi_t} - 1) B_{t-1}).
\end{aligned}
$$

Because of exogenous fiscal rules, the planner's instruments are: $a_{i,t}$, $l_{i,t}$, for individual variables and $\tilde{R}^{B,N}_t$, r_t, Π_t and ζ_t for aggregate variables.

Derivative with respect to $\tilde{R}^{B,N}_t$.

$$
\mathbb{E}_t \left[(1 - \tau^K_{t+1}) \frac{\Gamma_{t+1}}{\Pi_{t+1}} \right] B_t = \beta^{-1} \Upsilon_t \mathbb{E}_t \left[\frac{1}{\Pi_{t+1}} \right]. \tag{72}
$$

Derivative with respect to ζ_t.

$$
\begin{aligned}
0 = & w_t \int_i y_{i,t} \left(\psi_{i,t} l_{i,t} + \lambda_{i,l,t} u'(c_{i,t}) \right) \ell(di) - w_t \mu_t L_t \\
& + \frac{\varepsilon - 1}{\kappa} \zeta_t \gamma_t Y_t - (\tilde{r}^K_t + \delta) \left(\beta^{-1} \Upsilon_{t-1} + \Gamma_t (1 - \tau^K_t) K_{t-1} \right).
\end{aligned} \tag{73}
$$

Derivative with respect to Π_t.

$$0 = \mu_t \kappa (\Pi_t - 1) + (\gamma_t - \gamma_{t-1})(2\Pi_t - 1) + \left(\beta^{-1}\Upsilon_{t-1} - \Gamma_t(1-\tau_t^K)B_{t-1}\right) \frac{\tilde{R}_{t-1}^{B,N}}{Y_t \Pi_t^2}. \tag{74}$$

Derivative with respect to r_t.

$$0 = \int_i (\psi_{i,t} a_{i,t-1} + \lambda_{i,t-1} u'(c_{i,t}))\, \ell(di) + (\Gamma_t - \mu_t) A_{t-1}. \tag{75}$$

Derivative with respect to $a_{i,t}$. For unconstrained agents:

$$\begin{aligned}
\psi_{i,t} &= \beta \mathbb{E}_t \left[(1 + r_{t+1})\psi_{i,t+1}\right] \\
&+ \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[w_{t+1} \int_j y_{j,t+1} \left(\psi_{j,t+1} l_{j,t+1} + \lambda_{l,j,t+1} u'(c_{j,t+1})\right) \ell(dj) \right] \\
&- \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[\left((\gamma_{t+1} - \gamma_t)\Pi_{t+1}(\Pi_{t+1} - 1) - \frac{\varepsilon - 1}{\kappa} \gamma_{t+1}(\zeta_{t+1} - 1) \right) Y_{t+1} \right] \\
&+ \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[\mu_{t+1} \left(Y_{t+1} \left(1 - \frac{\kappa}{2}(\Pi_{t+1} - 1)^2 \right) - \frac{1}{\alpha}(r_{t+1} + \delta) K_t - w_{t+1} L_{t+1} \right) \right] \\
&- \beta \frac{\alpha - 1}{K_t} \mathbb{E}_t \left[(\tilde{r}_{t+1}^K + \delta) \left(\beta^{-1} \Upsilon_t + \Gamma_{t+1}(1 - \tau_{t+1}^K) K_t \right) \right] \\
&+ \beta \mathbb{E}_t \left[\Gamma_{t+1}(r_{t+1} - (1 - \tau_{t+1}^K)\tilde{r}_{t+1}^K) \right],
\end{aligned} \tag{76}$$

while for constrained ones we have $a_{i,t} = -\bar{a}$ and $\lambda_{i,t} = 0$.

Derivative with respect to $l_{i,t}$.

$$\begin{aligned}
\psi_{i,t} &= \frac{1}{w_t y_{i,t}} v'(l_{i,t}) + \frac{\lambda_{i,l,t}}{w_t y_{i,t}} v''(l_{i,t}) + \alpha \frac{1}{L_t} \int_j y_{j,t} \left(\psi_{j,t} l_{j,t} + \lambda_{l,j,t} u'(c_{j,t})\right) \ell(dj) \\
&+ \frac{1 - \alpha}{w_t L_t} \left[\mu_t w_t L_t + (\tilde{r}_t^K + \delta) \left(\beta^{-1} \Upsilon_{t-1} + \Gamma_t (1 - \tau_t^K) K_{t-1} \right) \right] \\
&+ \frac{1 - \alpha}{w_t L_t} \left((\gamma_t - \gamma_{t-1}) \Pi_t (\Pi_t - 1) - \frac{\varepsilon - 1}{\kappa} \gamma_t (\zeta_t - 1) - \mu_t \left(1 - \frac{\kappa}{2} (\Pi_t - 1)^2 \right) \right) Y_t
\end{aligned} \tag{77}$$

Derivative with respect to T_t.

$$\mu_t \leq \int_j \psi_{j,t} \ell(dj),$$

with equality if $T_t > 0$.

D The truncated representation

We derive the planner's first-order conditions for truncated economies.

D.1 The aggregation procedure

As explained in Section 2.6, the sequential representation characterizes the equilibrium of the full-fledged model by sequences of functions depending on the full aggregate and idiosyncratic histories of agents. Our aggregation procedure involves expressing the model based on so-called *truncated histories*, which are N-length vectors $h := (y_{-N+1}, \ldots, y_0)$ representing agents' idiosyncratic histories over the last N periods. The quantity $N \geq 0$ is the truncation length. Loosely speaking, we can represent the truncated history of an agent i whose idiosyncratic history is y^t as:

$$y^t = \{\ldots, y^h_{-N-2}, y^h_{-N-1}, y^h_{-N}, \underbrace{y^h_{-N+1}, \ldots, y^h_{-1}, y^h_0}_{=h}\}, \tag{78}$$

where h is the truncated history, corresponding to the idiosyncratic history over the last N periods. We now turn to the different steps of the aggregation procedure. We denote by $\mathcal{H} := \mathcal{Y}^N$ the set of all truncated histories.

First, the measure of agents with truncated history h, denoted by S_h, can be computed as:

$$S_h = \sum_{h \in \mathcal{H}} S_{\hat{h}} \pi_{\hat{h}h}, \tag{79}$$

where $\pi_{\hat{h}h}$ is the probability to switch from truncated history $\hat{h}$ at $t-1$ to truncated history h at t. It is equal to:

$$\pi_{\hat{h}h} = 1_{h \succeq \hat{h}} \pi_{y_0^{\hat{h}} y_0^h}, \tag{80}$$

and thus to the probability to switch from current productivity level $y_0^{\hat{h}}$ to current state y_0^h if h is a possible continuation of $\hat{h}$ (denoted by $h \succeq \hat{h}$).

Second, the model aggregation implies to assign consumption, saving, and labor choices to groups of agents with the same truncated history. For the sake of simplicity, we will write *truncated history* for "the group of agents sharing the same truncated history". Take the case of a generic variable, denoted by $X_t(y^t, z^t)$, where we make the dependence in y^t and z^t explicit. The quantity assigned to truncated history $h \in \mathcal{H}$ is denoted by $X_{t,h}$ and equal to the average value of X among agents with truncated history h. Formally:

$$X_{t,h} = \frac{1}{S_{t,h}} \sum_{y^t \in \mathcal{Y}^t \big| (y^t_{t-N+1}, \ldots, y^t_t) = h} X_t(y^t, z^t) \theta_t(y^t), \tag{81}$$

where $\theta_t(y^t)$ is recalled to be the measure of agents with history y^t. Definition (81) can be applied to the average consumption, the end-of-period saving, the labor supply, and the credit-constraint Lagrange multiplier respectively and lead to the quantities $c_{t,h}$, $a_{t,h}$, $l_{t,h}$, and $\nu_{t,h}$.

Third, we compute the aggregate beginning-of-period wealth. Applying definition (81) to period-t beginning-of-period wealth requires to account that agents transit from one truncated history at $t-1$ to another at t. The beginning-of-period wealth $\tilde{a}_{t,h}$ for truncated history h is:

$$\tilde{a}_{t,h} = \sum_{\hat{h}\in\mathcal{H}} \frac{S_{t-1,\hat{h}}}{S_{t,h}} \pi_{\hat{h}h} a_{t-1,\hat{h}}. \tag{82}$$

Fourth, the aggregation of the different equations characterizing the equilibrium is rather straightforward except for Euler equations – which involve non-linear marginal utilities. Indeed, the marginal utility of consumption aggregation ($u'(c_{t,h})$) is different from the aggregation of marginal utility ($u'_{t,h}$): $u'(c_{t,h}) \neq u'_{t,hj}$. The ratio of these two scalars will be denoted by $\xi_{t,h}$, which guarantees that Euler equations hold with aggregate consumption levels. Similarly, we denote by $\xi_{l,t,h}$ the parameters associated to the labor Euler equation.

D.2 The truncated model

We can now proceed with the aggregation of the full-fledged model. First, the aggregation of individual budget constraints (45) using equations (81) and (82) yields the following equation:

$$a_{t,h} + c_{t,h} = w_t y_0^h l_{t,h} + (1+r_t)\tilde{a}_{t,h} + T_t, \text{ for } h \in \mathcal{H}. \tag{83}$$

The aggregation of Euler equations for consumption (16) and labor (18) implies:

$$\xi_h u'(c_{t,h}) = \beta \mathbb{E}_t\left[(1+r_{t+1}) \sum_{\tilde{h}\succeq h} \pi_{h\tilde{h}} \xi_{\tilde{h}} u'(c_{t+1,\tilde{h}})\right] + \nu_{t,h}, \tag{84}$$

$$\xi_{l,h} v'(l_{t,h}) = w_t y_0^h \xi_h u'(c_{t,h}). \tag{85}$$

The system consisting of equations (83)–(85) is an exact aggregation of the full-fledged model with aggregate shocks in terms of truncated idiosyncratic histories. This system characterizes the dynamics of the aggregated variables $c_{t,h}$, $a_{t,h}$, $l_{t,h}$ and $\nu_{t,h}$ without any approximation.

Finally, we express market clearing conditions (19) using aggregated variables:

$$K_t = \sum_{h\in\mathcal{H}} S_{t,h} a_{t,h}, \quad L_t = \sum_{h\in\mathcal{H}} S_{t,h} y_0^h l_{t,h}. \tag{86}$$

The parameters ξs that appear in the aggregated Euler equations (84) are key in our method. We both show how to compute them using steady-state allocations and how these computations can be used to simulate the model in the presence of aggregate shocks.

Steady-state and computation of the ξs. At the steady state, the computations of the parameters ξs can be done based on allocations. Indeed, we can compute the stationary wealth distribution of the full-fledged model (using the individual equations) and identify credit-constrained histories. We can then compute aggregate (steady-state) allocations, c_h, a_h, l_h and ν_h, using equations (81) and (82). Then, consumption Euler equations (84) and (85) can be inverted to compute the parameters $(\xi_h)_h$ and $(\xi_{l,h})_h$ capturing the within-history heterogeneity in consumption and labor. Actually, this computation involves only standard linear algebra, and we provide a closed-form expression for the ξs – see equations (100) and (101) in Appendix D.5.

The truncated model in the presence of aggregate shocks. To use our truncation method in the presence of aggregate shocks, two further assumptions are needed.

Assumption A *We make the following two assumptions.*

1. *The parameters $(\xi_h)_h$ and $(\xi_{l,h})_h$ remain constant and equal to their steady-state values.*
2. *The set of credit-constrained histories, denoted by $\mathcal{C} \subset \mathcal{H}$, is time-invariant.*

The resulting model (i.e., the aggregated model with Assumption A) is called the truncated model. We therefore use the ξs twice: (i) once exactly to estimate them using the steady-state allocation; (ii) once approximately to simulate the model in the presence of aggregate shocks.[22]

Finally, two properties are worth mentioning. First, by construction of the ξs, the allocations of the full-fledged Bewley model and of the truncated equilibria coincide with each other at the steady state. Second, truncated equilibrium allocations (in the presence of aggregate shocks) can be proved to converge to those of the full-fledged equilibrium (and the ξs to 1), when the truncation length N becomes increasingly long. Furthermore, from a quantitative perspective, Section 5 shows that the ξs are an efficient tool to capture the heterogeneity within truncated histories, even when the truncation length is not too large.

D.3 Ramsey program

Thanks to its finite state-space representation, the truncated model makes it possible to solve the Ramsey program in the presence of aggregate shocks, which is a challenging task.[23] The Ramsey program in the truncated economy can be expressed as follows.

$$\max_{(w_t,r_t,\tilde{r}_t^K,\tilde{R}_t^N,\tau_t^K,\tau_t^L,B_t,T_t,K_t,L_t,\Pi_t,\zeta_t,(a_{t,h},c_{t,h},l_{t,h})_h)_{t\geq 0}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\sum_{h\in\mathcal{H}}S_h\xi_h(u(c_{t,h})-v(l_{t,h}))\right], \tag{87}$$

subject to aggregate Euler equations (84) and (85), aggregate budget constraint (83), aggregate market clearing conditions (86), credit constraints $a_{t,h} \geq -\bar{a}$, as well as constraints that were already present in the full-fledged Ramsey program: the governmental budget constraint (23),

[22] Assuming that the ξs are constant in the dynamics is equivalent to assuming that the distribution of agents within the truncated history is constant. This has the same spirit as assuming that the distribution within bins of wealth is uniform in the histogram method of Reiter (2009), among others.

[23] Note that we derive the first-order conditions of the truncated model, and we do not truncate the first-order conditions of the full-fledged Ramsey model. This ensures numerical stability, as the truncated model is "well-defined". It can be indeed considered as a model with limited insurance. See LeGrand and Ragot (2022a) for this expression and for the convergence of the FOCs of the truncated model toward the ones of the initial model.

the Phillips curve (30), the definition (2) of ζ_t, the one (4) of Y_t, the ones (9) and (10) of after-tax rates r_t, r_t^K, R_t^N and w_t, the zero profit condition for the fund (42), the no-arbitrage constraint (43), and the relationship (5) between factor prices.

As in Section 2.7, it is possible to use the tools of Marcet and Marimon (2019) to rewrite the Ramsey program. The truncation adds no complexity to the formulation of the planner's objective. First-order conditions can similarly be derived as in the general case, and we obviously have the same equivalence results.[24]

D.4 Program in the economy with sub-optimal fiscal policy

D.4.1 Program formulation

We take advantage of the equivalence result to simplify the program expression, which is:

$$\max_{(\tilde{R}_t^N, r_t, \zeta_t, B_t, (a_{t,h}, c_{t,h}, l_{t,h})_{h\in\mathcal{H}})_{t\geq 0}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{h\in\mathcal{H}} \Big[S_{t,h} \Big(u(c_{t,h}) - v(l_{t,h}) \tag{88}$$

$$- \left(\lambda_{t,h} - \tilde{\lambda}_{t,h}(1+r_t)\right) \xi_h u'(c_{t,h}) \Big) - \lambda_{l,t,h} \left(\xi_{l,h} v'(l_{t,h}) - w_t y_0^h \xi_h u'(c_{t,h}) \right) \Big) \Big],$$

$$\text{s.t. } c_{t,h} + a_{t,h} = w_t y_0^h l_{t,h} + (1+r_t)\tilde{a}_{t,h} + T_t, \tag{89}$$

$$a_{t,h} \geq 0 \text{ and } \tilde{a}_{t,h} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} \pi_{\tilde{h}h} \frac{S_{t-1,\tilde{h}}}{S_{t,h}} a_{t-1,\tilde{h}}, \tag{90}$$

$$\tilde{\lambda}_{t,h} = \frac{1}{S_{t,h}} \sum_{\tilde{h}\in\mathcal{H}} S_{t-1,\tilde{h}} \lambda_{t-1,\tilde{h}} \pi_{\tilde{h}h}, \tag{91}$$

$$K_t = \sum_h S_{t,h} a_{t,h} - B_t,\ L_t = \sum_h S_{t,h} y_h^0 l_{t,h}, \tag{92}$$

and subject to other constraints that are the same as in the individual case (government budget constraint, Phillips curve, fund no-arbitrage condition, fund zero profit condition, the definitions of ζ and of profits).

D.4.2 First-order conditions

We define the net social value of liquidity for history h similarly to definition (32):

$$\psi_{t,h} = \xi_h u'(c_{t,h}) - \left(\lambda_{t,h} - \tilde{\lambda}_{t,h}(1+r_t)\right) \xi_h u''(c_{t,h}) + \lambda_{l,h} y_0^h w_t u''(c_{t,h}). \tag{93}$$

The FOCs with respect to $\tilde{R}_t^N$, and Π_t are unchanged compared to the individual case.

[24] A final aspect regarding the truncated Ramsey program is that its solutions can be shown to converge to the solutions of the full-fledged Ramsey program (if they exist), when the truncation length N becomes infinitely long. See LeGrand and Ragot (2022a, Proposition 5). This convergence property is the parallel of the convergence result regarding allocations of the competitive equilibrium.

Derivative with respect to r_t.

$$0 = \sum_{h\in\mathcal{H}} S_{t,h}\psi_{t,h}\tilde{a}_{t,h} + \sum_{h\in\mathcal{H}} S_{t,h}\tilde{\lambda}_{t,h}\xi_h u'(c_{t,h}) + (\Gamma_t - \mu_t)A_{t-1}. \tag{94}$$

Derivative with respect to ζ_t.

$$\begin{aligned} 0 = {} & w_t \sum_{h\in\mathcal{H}} S_{t,h} y_0^h \left(l_{t,h}\psi_{t,h} + \lambda_{l,t,h}u'(c_{t,h})\right) - w_t\mu_t L_t \\ & + \frac{\varepsilon-1}{\kappa}\zeta_t\gamma_t Y_t - (\tilde{r}_t^K + \delta)\left(\beta^{-1}\Upsilon_{t-1} + \Gamma_t(1-\tau_t^K)K_{t-1}\right). \end{aligned}$$

Derivative with respect to $a_{t,h}$. For unconstrained truncated histories (i.e., $\nu_h = 0$):

$$\begin{aligned} \psi_{t,h} = {} & \beta\mathbb{E}_t\left[(1+r_{t+1})\sum_{\tilde{h}\in\mathcal{H}} \pi_{h\tilde{h}}\hat{\psi}_{t+1,\tilde{h}}\right] + \beta\mathbb{E}_t\left[\Gamma_{t+1}(r_{t+1} - (1-\tau_{t+1}^K)\tilde{r}_{t+1}^K)\right] \\ & + \beta\frac{\alpha}{K_t}\mathbb{E}_t\left[w_{t+1}\sum_{h\in\mathcal{H}} S_{t,h}y_0^h\left(l_{t+1,h}\psi_{t+1,h} + \lambda_{l,t+1,h}u'(c_{t+1,h})\right)\ell(dj)\right] \\ & - \beta\frac{\alpha}{K_t}\mathbb{E}_t\left[\left((\gamma_{t+1}-\gamma_t)\Pi_{t+1}(\Pi_{t+1}-1) - \frac{\varepsilon-1}{\kappa}\gamma_{t+1}(\zeta_{t+1}-1)\right)Y_{t+1}\right] \\ & + \beta\frac{\alpha}{K_t}\mathbb{E}_t\left[\mu_{t+1}\left(Y_{t+1}\left(1-\frac{\kappa}{2}(\Pi_{t+1}-1)^2\right) - \frac{1}{\alpha}(r_{t+1}+\delta)K_t - w_{t+1}L_{t+1}\right)\right] \\ & - \beta\frac{\alpha-1}{K_t}\mathbb{E}_t\left[(\tilde{r}_{t+1}^K + \delta)\left(\beta^{-1}\Upsilon_t + \Gamma_{t+1}(1-\tau_{t+1}^K)K_t\right)\right]. \end{aligned} \tag{95}$$

For constrained agents, we have $a_{t,h} + \bar{a} = 0$ and $\lambda_{t,h} = 0$.

Derivative with respect to $l_{t,h}$.

$$\begin{aligned} \psi_{t,h} = {} & \frac{1}{w_t y_0^h}v'(l_{t,h}) + \frac{\lambda_{l,t,h}}{w_t y_0^h}v''(l_{t,h}) + \alpha\frac{1}{L_t}\sum_{h\in\mathcal{H}} S_{t,h}y_0^h\left(l_{t,h}\psi_{t,h} + \lambda_{l,t,h}u'(c_{t,h})\right) \\ & + \frac{1-\alpha}{w_t L_t}\left(\mu_t w_t L_t + (\tilde{r}_t^K + \delta)\left(\beta^{-1}\Upsilon_{t-1} + \Gamma_t(1-\tau_t^K)K_{t-1}\right)\right) \\ & + \frac{1-\alpha}{w_t L_t}\left((\gamma_t - \gamma_{t-1})\Pi_t(\Pi_t - 1) - \frac{\varepsilon-1}{\kappa}\gamma_t(\zeta_t - 1) - \mu_t\left(1 - \frac{\kappa}{2}(\Pi_t-1)^2\right)\right)Y_t. \end{aligned} \tag{96}$$

Derivative with respect to T_t.

$$\sum_h S_{t,h}\psi_{t,h} \leq \mu_t, \tag{97}$$

with equality when $T_t > 0$.

D.5 Matrix representation of the steady-state allocation

At the steady-state, we have $Z = 1$, and the FOCs (72) and (74) (with respect to $\tilde{R}_t^N$ and Π_t) imply that $\Pi = 1$, and using the Phillips curve, $\zeta = 1$.[25]

In this section, we provide closed-form formulas for parameters ξs and the Lagrange multiplier, so as to solve the steady-state model. We will denote by $\circ$ the Hadamard product and by $\times$ or an absence of mathematical sign the usual matrix product.

The matrix representation consists in stacking together the equations characterizing the steady-state, so as to provide a convenient matrix notation for solving the steady state. It also provides an efficient solution to compute the values for the ξs and the Lagrange multiplier. We recall that there are N_{tot} truncated histories and we assume that the set of truncated histories is endowed with a total order that is used for indexing vectors and matrices. In other words, $(x_h)_h$ implicitly assumes that all elements x_h are collected along a given order that remains identical from one vector to another.

Let $\boldsymbol{S}$ be the N_{tot}-vector of steady-state history sizes that is defined as $\boldsymbol{S} = (S_h)_h$. Similarly, let $\boldsymbol{a}$, $\boldsymbol{c}$, and $\boldsymbol{\nu}$ be the N_{tot}-vectors of end-of-period wealth, consumption, credit-constraint Lagrange multipliers, respectively. These vectors are known from the steady-state equilibrium of the Bewley model. Denote by $\boldsymbol{I}$ the $(N_{tot} \times N_{tot})$-identity matrix, $\mathbf{1}_{N_{tot}}$ the $(N_{tot} \times 1)$-vector of 1, and $\boldsymbol{\Pi}$ as the transition matrix across histories. We obtain the following steady-state relationships:

$$\boldsymbol{S} = \boldsymbol{\Pi}^\top \boldsymbol{S}, \tag{98}$$

$$\boldsymbol{S} \circ \boldsymbol{c} + \boldsymbol{S} \circ \boldsymbol{a} = (1+r)\boldsymbol{\Pi}^\top (\boldsymbol{S} \circ \boldsymbol{a}) + \boldsymbol{S} \circ \boldsymbol{W} \circ \boldsymbol{l}, \tag{99}$$

where (98) is the dynamics of history sizes and (99) is the history of budget constraints.

Euler equations (84) and (85) imply that we are looking for vectors $\boldsymbol{\xi}^u$ and $\boldsymbol{\xi}^v$ such that:

$$\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) = \beta(1+r)\boldsymbol{\Pi}\left(\boldsymbol{\xi}^u \circ u'(\boldsymbol{c})\right) + \boldsymbol{\nu}, \tag{100}$$

$$\boldsymbol{\xi}^v \circ v'(\boldsymbol{l}) = w\boldsymbol{\xi}^u \circ \boldsymbol{y} \circ u'(\boldsymbol{c}), \tag{101}$$

where the matrix $\boldsymbol{\Pi}^\top$ (the transpose of $\boldsymbol{\Pi}$) is used to make expectations about next-period histories and where $u'(\boldsymbol{c})$ assumes that the function applies element-wise (and thus denotes $(u'(c_h))_h$). Finally, for any vector $\boldsymbol{x}$, we denote by $\boldsymbol{D_x}$ the diagonal matrix with $\boldsymbol{x}$ on the diagonal. We deduce from (100) that $\boldsymbol{\xi}^u$ is defined by the following closed-form formula:

$$u'(\boldsymbol{c}) \circ \boldsymbol{\xi} = (\boldsymbol{I} - \beta(1+r)\boldsymbol{\Pi})^{-1}\boldsymbol{\nu}. \tag{102}$$

We then use $\boldsymbol{\xi}^v \circ v'(\boldsymbol{l}) = w\boldsymbol{\xi}^u \circ \boldsymbol{y} \circ u'(\boldsymbol{c})$ to compute $\boldsymbol{\xi}^v$.

We set an initial guess value for μ. We start with the definitions of the $\tilde{\lambda}_c$s and of the ψs of

[25] We remove the t subscript to denote steady-state quantities.

equations (91) and (93):

$$\boldsymbol{S} \circ \tilde{\boldsymbol{\lambda}}_c = \boldsymbol{\Pi}^\top (\boldsymbol{S} \circ \boldsymbol{\lambda}_c), \tag{103}$$

$$\boldsymbol{\psi} = \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) - \boldsymbol{D}_{\boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} \left(\boldsymbol{\lambda}_c - (1+r)\tilde{\boldsymbol{\lambda}}_c - w\boldsymbol{y} \circ \boldsymbol{\lambda}_l \right). \tag{104}$$

Combining these two equations yields:

$$\boldsymbol{S} \circ \boldsymbol{\psi} = \boldsymbol{S} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) - \boldsymbol{D}_{\boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} \left(\boldsymbol{I} - (1+r)\boldsymbol{\Pi}^\top \right) (\boldsymbol{S} \circ \boldsymbol{\lambda}_c) + \boldsymbol{D}_{w\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} (\boldsymbol{S} \circ \boldsymbol{\lambda}_l), \tag{105}$$

where $\circ$ is commutative and for any vectors $\boldsymbol{x}$ and $\boldsymbol{y}$: $\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{D}_{\boldsymbol{x}} \boldsymbol{y}$ and $\boldsymbol{D}_{\boldsymbol{x}} \boldsymbol{D}_{\boldsymbol{y}} = \boldsymbol{D}_{\boldsymbol{x} \circ \boldsymbol{y}}$.

We also denote: $\tilde{\xi}^v_h = \frac{\xi^v_h}{y^h_0}$. The FOC (96) for labor becomes:

$$\begin{aligned}
\boldsymbol{S} \circ \boldsymbol{\psi} &= \frac{1}{w} \boldsymbol{S} \circ \tilde{\boldsymbol{\xi}}^v \circ v'(\boldsymbol{l}) + \frac{1}{w} \tilde{\boldsymbol{\xi}}^v \circ v''(\boldsymbol{l}) \circ (\boldsymbol{S} \circ \boldsymbol{\lambda}_l) \\
&+ (\alpha - 1) \frac{r + (1-\tau^K)\delta}{wL} \left(\boldsymbol{S} \tilde{\boldsymbol{a}}^\top (\boldsymbol{S} \circ \boldsymbol{\psi}) + \boldsymbol{S} (\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top (\boldsymbol{S} \circ \tilde{\boldsymbol{\lambda}}_c) \right) \\
&+ \alpha \frac{1}{L} \left(\boldsymbol{S} (\boldsymbol{y} \circ \boldsymbol{l})^\top (\boldsymbol{S} \circ \boldsymbol{\psi}) + \boldsymbol{S} (\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top (\boldsymbol{S} \circ \tilde{\boldsymbol{\lambda}}_l) \right) \\
&- \frac{\mu}{wL} (1-\alpha) \left(Y - (r + (1-\tau^K)\delta) A - wL \right) \boldsymbol{S}.
\end{aligned}$$

Denoting $\boldsymbol{M}_0 = wL\boldsymbol{I} - (\alpha - 1)(r + (1-\tau^K)\delta) \boldsymbol{S} \tilde{\boldsymbol{a}}^\top - \alpha w \boldsymbol{S} (\boldsymbol{y} \circ \boldsymbol{l})^\top$, we have:

$$\begin{aligned}
\boldsymbol{M}_0 \boldsymbol{S} \circ \boldsymbol{\psi} &= (\alpha - 1)(r + (1-\tau^K)\delta)(\boldsymbol{S} (\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi}^\top)(\boldsymbol{S} \circ \boldsymbol{\lambda}_c) \\
&+ (\boldsymbol{D}_{L\tilde{\boldsymbol{\xi}}^v \circ v''(\boldsymbol{l})} + \alpha w (\boldsymbol{S} (\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi}^\top))(\boldsymbol{S} \circ \boldsymbol{\lambda}_l) \\
&+ L\boldsymbol{S} \circ \boldsymbol{\omega} \circ \tilde{\boldsymbol{\xi}}^v \circ v'(\boldsymbol{l}) - \mu(1-\alpha) \left(Y - (r + (1-\tau^K)\delta) A - wL \right) \boldsymbol{S},
\end{aligned} \tag{106}$$

where it should be observed that $\boldsymbol{S}(\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi}^\top$ has the dimension $(N_{tot} \times 1) \cdot (1 \times N_{tot}) \cdot (N_{tot} \times N_{tot}) = (N_{tot} \times N_{tot})$ and is a matrix. Similarly, $\boldsymbol{S}(\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) \circ \boldsymbol{y})^\top \boldsymbol{\Pi}^\top$, $\boldsymbol{S}\tilde{\boldsymbol{a}}^\top$, and $\boldsymbol{S}(\boldsymbol{y} \circ \boldsymbol{l})^\top$ also have the dimension $(N_{tot} \times 1) \cdot (1 \times N_{tot}) = (N_{tot} \times N_{tot})$ and are matrices.

Using (105):

$$\boldsymbol{S} \circ \boldsymbol{\psi} = \boldsymbol{\omega} \circ \boldsymbol{S} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) - \boldsymbol{D}_{\boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} \left(\boldsymbol{I} - (1+r)\boldsymbol{\Pi}^\top \right) (\boldsymbol{S} \circ \boldsymbol{\lambda}_c) + \boldsymbol{D}_{w\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} (\boldsymbol{S} \circ \boldsymbol{\lambda}_l),$$

we deduce:

$$\begin{aligned}
&\boldsymbol{M}_0 \boldsymbol{\omega} \circ \boldsymbol{S} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}) - L\boldsymbol{S} \circ \boldsymbol{\omega} \circ \tilde{\boldsymbol{\xi}}^v \circ v'(\boldsymbol{l}) + \mu(1-\alpha) \left(Y - (r + (1-\tau^K)\delta) A - wL \right) \boldsymbol{S} \\
&= \left(\boldsymbol{M}_0 \boldsymbol{D}_{\boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} \left(\boldsymbol{I} - (1+r)\boldsymbol{\Pi}^\top \right) + (\alpha - 1)(r + (1-\tau^K)\delta) \boldsymbol{S} (\boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi}^\top \right) (\boldsymbol{S} \circ \boldsymbol{\lambda}_c) \\
&+ (-\boldsymbol{M}_0 \boldsymbol{D}_{w\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u''(\boldsymbol{c})} + \boldsymbol{D}_{L\tilde{\boldsymbol{\xi}}^v \circ v''(\boldsymbol{l})} + \alpha w \boldsymbol{S} (\boldsymbol{y} \circ \boldsymbol{\xi}^u \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi}^\top)(\boldsymbol{S} \circ \boldsymbol{\lambda}_l),
\end{aligned}$$

or:

$$\boldsymbol{S}\circ\boldsymbol{\lambda}_c = \boldsymbol{L}_0(\boldsymbol{S}\circ\boldsymbol{\lambda}_l)+\boldsymbol{x}_0, \tag{107}$$

$$\text{where: } \tilde{\boldsymbol{L}}_0 = \boldsymbol{M}_0\boldsymbol{D}_{\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}\left(\boldsymbol{I}-(1+r)\boldsymbol{\Pi}^\top\right)+(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{S}(\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top,$$

$$\boldsymbol{L}_0 = \tilde{\boldsymbol{L}}_0^{-1}(\boldsymbol{M}_0\boldsymbol{D}_{w\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}-\boldsymbol{D}_{L\tilde{\boldsymbol{\xi}}^v\circ v''(\boldsymbol{l})}-\alpha w(\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top)),$$

$$\boldsymbol{x}_0 = \tilde{\boldsymbol{L}}_0^{-1}\Big(\boldsymbol{M}_0\boldsymbol{\omega}\circ\boldsymbol{S}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c})-L\boldsymbol{S}\circ\boldsymbol{\omega}\circ\tilde{\boldsymbol{\xi}}^v\circ v'(\boldsymbol{l})$$
$$+\mu(1-\alpha)\left(Y-(r+(1-\tau^K)\delta)A-wL\right)\boldsymbol{S}\Big).$$

We now turn to the Euler-like equation (95) that holds only for unconstrained histories. We introduce the matrix $\boldsymbol{\Pi}^\psi$ defined for any vector $\boldsymbol{x}$ as:

$$\boldsymbol{\Pi}^\psi(\boldsymbol{S}\circ\boldsymbol{x}) = \boldsymbol{S}\circ(\boldsymbol{\Pi}\boldsymbol{x}), \tag{108}$$

and the matrix $\boldsymbol{P}$, which is the matrix with one on the diagonal if the history is unconstrained and 0 otherwise. With this notation, equation (95) becomes:

$$\begin{aligned}\boldsymbol{P}(\boldsymbol{S}\circ\boldsymbol{\psi}) &= \beta(1+r)\boldsymbol{P}\boldsymbol{\Pi}^\psi(\boldsymbol{S}\circ\boldsymbol{\psi}) \\ &+\frac{\beta}{K}(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{P}\left(\boldsymbol{S}\tilde{\boldsymbol{a}}^\top(\boldsymbol{S}\circ\boldsymbol{\psi})+\boldsymbol{S}(\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top(\boldsymbol{S}\circ\tilde{\boldsymbol{\lambda}}_c)\right) \\ &+\frac{\beta}{K}\alpha w\boldsymbol{P}\left(\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{l})^\top(\boldsymbol{S}\circ\boldsymbol{\psi})+\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top(\boldsymbol{S}\circ\tilde{\boldsymbol{\lambda}}_l)\right) \\ &+\frac{\beta}{K}\boldsymbol{P}\mu\left(\alpha Y-(\alpha-1)(r+(1-\tau^K)\delta)A-(r+\delta)K-\alpha wL\right)\boldsymbol{S},\end{aligned} \tag{109}$$

Denoting: $\tilde{\boldsymbol{L}}_1 = \boldsymbol{I}-\beta(1+r)\boldsymbol{\Pi}^\psi-\frac{\beta}{K}(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{S}\tilde{\boldsymbol{a}}^\top-\frac{\beta}{K}\alpha w\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{l})^\top$, equation (109) becomes:

$$\begin{aligned}\boldsymbol{P}\tilde{\boldsymbol{L}}_1(\boldsymbol{S}\circ\boldsymbol{\psi}) &= \frac{\beta}{K}(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{P}\boldsymbol{S}(\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top(\boldsymbol{S}\circ\boldsymbol{\lambda}_c) \\ &+\frac{\beta}{K}\alpha w\boldsymbol{P}\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top(\boldsymbol{S}\circ\boldsymbol{\lambda}_l) \\ &+\frac{\beta}{K}\boldsymbol{P}\mu\left(\alpha Y-(\alpha-1)(r+(1-\tau^K)\delta)A-(r+\delta)K-\alpha wL\right)\boldsymbol{S}.\end{aligned}$$

Using (105) to express $\boldsymbol{S}\circ\boldsymbol{\psi}$, we obtain:

$$\begin{aligned}&\boldsymbol{P}\left(\tilde{\boldsymbol{L}}_1\boldsymbol{D}_{\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}\left(\boldsymbol{I}-(1+r)\boldsymbol{\Pi}^\top\right)+\frac{\beta}{K}(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{S}(\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top\right)(\boldsymbol{S}\circ\boldsymbol{\lambda}_c) \\ &=\boldsymbol{P}\tilde{\boldsymbol{L}}_1\boldsymbol{\omega}\circ\boldsymbol{S}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c})-\frac{\beta}{K}\boldsymbol{P}\mu\left(\alpha Y-(\alpha-1)(r+(1-\tau^K)\delta)A-(r+\delta)K-\alpha wL\right)\boldsymbol{S} \\ &+\boldsymbol{P}\left(\tilde{\boldsymbol{L}}_1\boldsymbol{D}_{w\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}-\frac{\beta}{K}\alpha w\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top\right)(\boldsymbol{S}\circ\boldsymbol{\lambda}_l).\end{aligned}$$

We finally deduce:

$$\boldsymbol{P}\boldsymbol{L}_{1,c}(\boldsymbol{S}\circ\boldsymbol{\lambda}_c) = \boldsymbol{P}\boldsymbol{L}_{1,l}(\boldsymbol{S}\circ\boldsymbol{\lambda}_l) + \boldsymbol{P}\boldsymbol{x}_1, \tag{110}$$

$$\begin{aligned}
\text{where: } \tilde{\boldsymbol{L}}_1 &= \boldsymbol{I} - \beta(1+r)\boldsymbol{\Pi}^{\psi} - \frac{\beta}{K}(\alpha-1)(r+(1-\tau^K)\delta)\boldsymbol{S}\tilde{\boldsymbol{a}}^\top - \frac{\beta}{K}\alpha w\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{l})^\top,\\
\boldsymbol{L}_{1,c} &= \tilde{\boldsymbol{L}}_1\boldsymbol{D}_{\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}\left(\boldsymbol{I}-(1+r)\boldsymbol{\Pi}^\top\right) + \frac{\beta}{K}(\alpha-1)(r+\delta)\boldsymbol{S}(\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top,\\
\boldsymbol{L}_{1,l} &= \tilde{\boldsymbol{L}}_1\boldsymbol{D}_{w\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})} - \frac{\beta}{K}\alpha w\boldsymbol{S}(\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}^\top,\\
\boldsymbol{x}_1 &= \tilde{\boldsymbol{L}}_1\boldsymbol{\omega}\circ\boldsymbol{S}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c})\\
&\quad - \frac{\beta}{K}\mu\left(\alpha Y - (\alpha-1)(r+(1-\tau^K)\delta)A - (r+\delta)K - \alpha wL\right)\boldsymbol{S}.
\end{aligned}$$

We also have for unconstrained agents: $(\boldsymbol{I}-\boldsymbol{P})(\boldsymbol{S}\circ\boldsymbol{\lambda}_c)=0$, which yields with (110):

$$\begin{aligned}
\boldsymbol{S}\circ\boldsymbol{\lambda}_c &= \boldsymbol{L}_2(\boldsymbol{S}\circ\boldsymbol{\lambda}_l) + \boldsymbol{x}_2, \qquad (111)\\
\text{where: } \boldsymbol{L}_2 &= (\boldsymbol{I}-\boldsymbol{P}+\boldsymbol{P}\boldsymbol{L}_{1,c})^{-1}\boldsymbol{P}\boldsymbol{L}_{1,l},\\
\boldsymbol{x}_2 &= (\boldsymbol{I}-\boldsymbol{P}+\boldsymbol{P}\boldsymbol{L}_{1,c})^{-1}\boldsymbol{P}\boldsymbol{x}_1.
\end{aligned}$$

Combining (107) and (111), we obtain:

$$\begin{aligned}
\boldsymbol{S}\circ\boldsymbol{\lambda}_l &= -(\boldsymbol{L}_0-\boldsymbol{L}_2)^{-1}(\boldsymbol{x}_0-\boldsymbol{x}_2),\\
\boldsymbol{S}\circ\boldsymbol{\lambda}_c &= \boldsymbol{x}_2 - \boldsymbol{L}_2(\boldsymbol{L}_2-\boldsymbol{L}_0)^{-1}(\boldsymbol{x}_2-\boldsymbol{x}_0),
\end{aligned}$$

and we deduce:

$$\begin{aligned}
\boldsymbol{S}\circ\boldsymbol{\psi} &= \boldsymbol{S}\circ\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}) - \boldsymbol{D}_{\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}\left(\boldsymbol{I}-(1+r)\boldsymbol{\Pi}^\top\right)\boldsymbol{x}_2 + \boldsymbol{D}_{w\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}(\boldsymbol{S}\circ\boldsymbol{\lambda}_l), \qquad (112)\\
&\quad + (\boldsymbol{D}_{\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})}\left(\boldsymbol{I}-(1+r)\boldsymbol{\Pi}^\top\right) - \boldsymbol{D}_{w\boldsymbol{y}\circ\boldsymbol{\xi}^u\circ u''(\boldsymbol{c})})(\boldsymbol{L}_0-\boldsymbol{L}_2)^{-1}(\boldsymbol{x}_0-\boldsymbol{x}_2).
\end{aligned}$$

It remains to iterate on μ until FOC (97) holds, or: $\mu = \mathbf{1}^\top(\boldsymbol{S}\circ\boldsymbol{\psi})$.

We then compute the steady-state value of other Lagrange multipliers using the FOCs (72), (73), and (94):

$$\begin{aligned}
\Gamma &= \mu - \frac{1}{\boldsymbol{S}^\top\boldsymbol{a}}\left(\tilde{\boldsymbol{a}}^\top(\boldsymbol{S}\circ\boldsymbol{\psi}) + (\boldsymbol{\xi}^u\circ u'(\boldsymbol{c}))^\top(\boldsymbol{S}\circ\boldsymbol{\lambda}_c)\right),\\
\Upsilon &= \beta(1-\tau^K)\Gamma B,\\
\frac{\varepsilon-1}{\kappa}\gamma Y &= (\tilde{r}^K+\delta)\left(\beta^{-1}\Upsilon + \Gamma_t(1-\tau^K)K\right) + w\mu L\\
&\quad - w(\boldsymbol{y}\circ\boldsymbol{l})^\top(\boldsymbol{S}\circ\boldsymbol{\psi}) - w(\boldsymbol{y}\circ u'(\boldsymbol{c}))^\top(\boldsymbol{S}\circ\boldsymbol{\lambda}_l).
\end{aligned}$$

E A robustness check: The refined truncation

E.1 The refinement method

The refinement method, developed in LeGrand and Ragot (2022c), consists in considering truncated histories of unequal length, instead of truncated histories of the same length N, as in Section D. The idea is that some histories that are particularly large or heterogeneous would benefit from longer length. For example, with two idiosyncratic states (corresponding to high productivity y_2 and to low productivity y_1), the set of truncated histories of length 2 is $\{(y_2, y_2), (y_2, y_1), (y_1, y_2), (y_1, y_1)\}$, where (y_2, y_2) and (y_1, y_1) are typically the two largest truncated histories, since productivity is persistent. In that case, the refinement method allows one to split these larger histories, without affecting the shorter histories (y_2, y_1), and (y_1, y_2). For instance, the history (y_2, y_2) can be refined into (y_2, y_2, y_2) and (y_1, y_2, y_2), such that the set of truncated histories is now $\{(y_2, y_2, y_2), (y_1, y_2, y_2), (y_2, y_1), (y_1, y_2), (y_1, y_1)\}$. It can be checked that: (i) this set corresponds to a proper partition of idiosyncratic histories (every infinite history admits a unique truncated representation in the set) and (ii) the transition matrix between set histories is well-defined.[26]

With k idiosyncratic states, the refinement method works as follows. We denote a set of refined truncated histories by $R(N, (N_i)_{i=1,\dots,k})$, where N is the uniform truncation length (on which the refinement is based), and $N_i \geq N$ is the longest refinement history for the state i. The construction is recursive and follows the same steps as the above example. It starts from the set $R(N, (N)_{i=1,\dots,k})$ of uniform truncated histories that are all of length N. The first refinement step consists in substituting for the history $y_1^N = (\underbrace{y_1, \dots, y_1}_{N})$, the set of histories $\{(y_i, \underbrace{y_1, \dots, y_1}_{N}) : i = 1, \dots, k\}$. This yields the set $R(N, N+1, (N_i)_{i=2,\dots,k})$ of size $k^N + k - 1$. The second step consists in refining the history $y_1^{N+1} = (\underbrace{y_1, \dots, y_1}_{N+1})$ into $\{(y_i, \underbrace{y_1, \dots, y_1}_{N+1}) : i = 1, \dots, k\}$, which gives the set $R(N, N+2, (N_i)_{i=2,\dots,k})$ – of size $k^N + 2(k-1)$. These steps are repeated until we obtain $R(N, N_1, (N_i)_{i=2,\dots,k})$, which is of size $k^N + (N_1 - N)(k-1)$. Furthermore, this set implies a well-defined mapping between infinite and truncated histories, as well as a well-defined transition probability matrix. The same method can be applied to any other history of the form $y_i^N = (\underbrace{y_i, \dots, y_i}_{N})$ with $i = 2, \dots, k$. This finally generates the set $R(N, (N_i)_{i=1,\dots,k})$, with $k^N + (k-1)\sum_{i=1}^{k}(N_i - N)$ elements.

The truncation method can then be used as in the uniform case of Section D. Equations characterizing the Ramsey allocation in Section D.4 and the matrix representation of Section D.5 remain valid once the proper transition matrix has been computed (see equation (80)).

[26]The last point does not hold in general. Indeed, $\{(y_2, y_2), (y_2, y_1, y_2), (y_1, y_1, y_2), (y_1)\}$ corresponds to the refinement of $\{(y_2, y_2), (y_1, y_2), (y_1)\}$ (through (y_1, y_2)) but does not yield a well-defined transition matrix.

E.2 Quantitative application to the model of Section 5

We consider a refined truncation based on a uniform truncation with $N = 5$. We then consider a common refinement length of 20 periods for all constant histories (of the form $y_i^N = (y_i, \ldots, y_i)$). The number of truncated histories with positive size increases to 907 from 727 with the uniform truncation method. We report in Table 5 the second-order moments associated to the simulation of the benchmark economy and the economy with time-varying taxes using the refined truncation method. As can be seen, the differences with the uniform truncation method are small.

		Benchmark	Time-varying taxes
Y	Mean	1.43	1.43
	Std(%)	1.48	1.47
C	Mean	0.9	0.9
	Std(%)	1.33	1.31
K	Mean	14.31	14.31
	Std(%)	1.54	1.60
L	Mean	0.39	0.39
	Std(%)	0.28	0.18
τ^L	Mean	0.28	0.28
	Std(%)	0.0	0.2
τ^K	Mean	0.36	0.36
	Std(%)	0.0	0.2
B	Mean	3.64	3.64
	Std(%)	0.62	0.62
T	Mean	0.11	0.11
	Std(%)	4.56	5.70
Π	Mean	1.00	1.00
	Std(%)	0.026	0.017
	Correlations		
$corr(\Pi, Y)$		0.22	0.20
$corr(\tau^K, Y)$		0.0	-0.52
$corr(\tau^L, Y)$		0.0	0.52
$corr(B, Y)$		-0.97	-0.97
$corr(C, Y)$		0.95	0.95
$corr(Y, Y_{-1})$		0.98	0.99
$corr(B, B_{-1})$		0.96	0.96

Table 5: Refined truncation (20 periods). First- and second-order moments for key variables, in the three economies (Economy 1 is the benchmark with Fiscal Rule 1; Economy 2 is the economy with constant inflation rate; Economy 3 is the economy with Fiscal Rule 2).

F Comparisons with the Reiter method

Since Section E.2 has shown that the uniform and refined truncations yield similar results, we compare the Reiter method to the refined truncation method of Section E.2 (with $N = 5$ and a refinement length of 20). We assume full price flexibility and no public spending need ($G = 0$), and hence no tax. Table 6 presents the calibration and we still use 5 idiosyncratic states.

Parameter	Description	Value
	Preference and technology	
β	Discount factor	0.99
σ	Curvature utility	1
α	Capital share	0.36
δ	Depreciation rate	0.025
$\bar{a}$	Credit limit	0
φ	Frisch elasticity labor supply	0.5
	Shock process	
ρ_z	Autocorrelation TFP	0.95
σ_z	Standard deviation TFP shock	0.31%
ρ_y	Autocorrelation idio. income	0.99
σ_y	Standard dev. idio. income	12%

Table 6: Parameter values in the baseline calibration. See text for descriptions and targets.

First, Figure 5 plots the IRFs for the main variables after a negative TFP shock, calibrated to 1%. The two models are very close, and almost indistinguishable. Second, we report first and

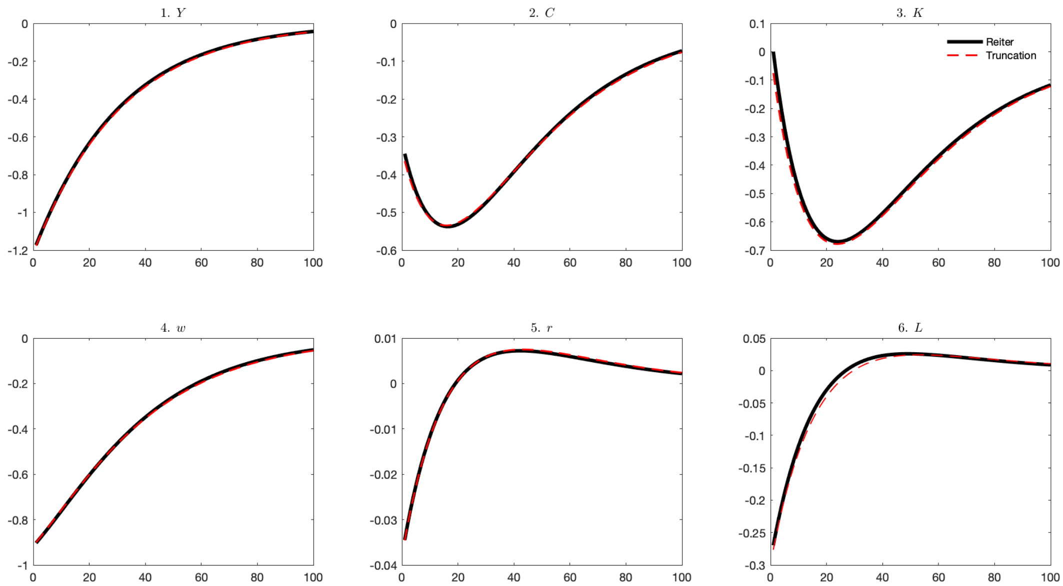


Figure 5: IRFs after a negative TFP shock, in percentage deviation for all variables, except r in level deviation. The black solid line is Reiter method and red dashed line is truncation method.

second-order moments. First-order moments are similar by construction, as the two methods involve a perturbation method around the same steady state. Normalized standard deviations are reported in percent to ease the reading. Again, second-order moments are very close. The difference is around 10^{-4}, which is a good outcome.

		Reiter	Truncation
Y	Mean	1.4772	1.4772
	Std(%)	1.4365	1.4521
C	Mean	1.0252	1.0252
	Std(%)	1.0554	1.0593
K	Mean	18.0831	18.0831
	Std(%)	1.3338	1.3667
L	Mean	0.3610	0.3610
	Std(%)	0.2105	0.2216
r	Mean	0.0044	0.0044
	Std(%)	0.0280	0.0286
w	Mean	2.6187	2.6187
	Std(%)	1.2927	1.2930
	Correlations		
$corr(C, Y)$		0.9064	0.9130
$corr(Y, Y_{-1})$		0.9668	0.9674

Table 7: First- and second-order moments for key variables, comparing the same model simulated with the Reiter's method and the refined truncation method.

Third, we simulate the model for 10,000 periods to report the average and maximum absolute normalized differences (i.e. divided by the mean of the variables) for Y, C, K, L, w and absolute normalized difference for r between the two simulation methods.

	Y	C	K	L	r	w
Average	$1.60 \cdot 10^{-4}$	$1.37 \cdot 10^{-4}$	$4.73 \cdot 10^{-4}$	$1.36 \cdot 10^{-4}$	$6.37 \cdot 10^{-6}$	$1.22 \cdot 10^{-4}$
Maximum	$8.34 \cdot 10^{-4}$	$6.54 \cdot 10^{-4}$	$2.44 \cdot 10^{-3}$	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-5}$	$6.72 \cdot 10^{-4}$

Table 8: Average and maximum absolute differences for key variables between the two methods for simulation over 10,000 periods.

G Unequal profit distribution

In this section, we investigate the case of an unequal profit distribution. This corresponds to the setup of Section 4, but in which the firms' profits are unequally distributed to households instead of being fully taxed by the government. A s explained in Section 6.2, agents of type i receive a profit amount equal to $(\sum S_y y)^{-1} y_i^\nu \Omega_t$, where $\nu \geq 0$ drives how unequal the profit distribution is.

G.1 The Ramsey program

The Ramsey program is similar to the one of Section 4.3, except that the budget constraints of households and of the government (equations (45) and (48)) are modified as follows:

$$a_{i,t} + c_{i,t} = (1+r_t)a_{i,t-1} + w_t y_{i,t} l_{i,t} + T_t + \frac{y_{i,t}^\nu}{\sum_y S_y y^\nu}\left(1 - \zeta_t - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Y_t,$$

$$B_t + \zeta_t Y_t - \delta K_{t-1} = G_t + B_{t-1} + r_t A_{t-1} + w_t L_t + T_t.$$

To save space, we directly present the FOCs of the truncated model. The social valuation of liquidity $\psi_{t,h}$ is still defined in equation (93).

FOCs with respect to $\tilde{R}_t^{B,N}$, r_t, and T_t. These FOCs are unchanged compared to the full profit taxation of Section 4.3 and are grouped below:

$$\mathbb{E}_t\left[(1-\tau_{t+1}^K)\frac{\Gamma_{t+1}}{\Pi_{t+1}}\right] B_t = \beta^{-1}\Upsilon_t \mathbb{E}_t\left[\frac{1}{\Pi_{t+1}}\right], \tag{113}$$

$$(\mu_t - \Gamma_t)A_{t-1} = \sum_{h\in\mathcal{H}} S_{t,h}\left(\psi_{t,h}\tilde{a}_{t,h} + \tilde{\lambda}_{t,h}\xi_h u'(c_{t,h})\right), \tag{114}$$

$$\mu_t = \sum_{h\in\mathcal{H}} S_{t,h}\psi_{t,h}. \tag{115}$$

The other FOCs are modified with new profit distribution.

FOC with respect to ζ_t.

$$\begin{aligned} 0 = {} & w_t \sum_{h\in\mathcal{H}} S_{t,h} y_0^h \left(l_{t,h}\psi_{t,h} + \lambda_{l,t,h} u'(c_{t,h})\right) - w_t \mu_t L_t \\ & + \left(\frac{\varepsilon - 1}{\kappa}\gamma_t + \mu_t - \frac{1}{\sum_y S_y y^\nu}\sum_{h\in\mathcal{H}} S_{t,h}(y_0^h)^\nu \psi_{t,h}\right)\zeta_t Y_t \\ & - (\tilde{r}_t^K + \delta)\left(\beta^{-1}\Upsilon_{t-1} + \Gamma_t(1-\tau_t^K)K_{t-1}\right). \end{aligned} \tag{116}$$

FOC with respect to Π_t.

$$\begin{aligned} 0 = {} & \frac{\kappa(\Pi_t - 1)}{\sum_y S_y y^\nu}\sum_{h\in\mathcal{H}} S_{t,h}(y_0^h)^\nu \psi_{t,h} + (\gamma_t - \gamma_{t-1})(2\Pi_t - 1) \\ & + \left(\beta^{-1}\Upsilon_{t-1} - \Gamma_t(1-\tau_t^K)B_{t-1}\right)\frac{\tilde{R}_{t-1}^{B,N}}{Y_t \Pi_t^2}. \end{aligned} \tag{117}$$

FOC with respect to $a_{t,h}$. For unconstrained agents:

$$\begin{aligned}
\psi_{t,h} = {} & \beta \mathbb{E}_t \left[(1+r_{t+1}) \sum_{\tilde{h} \in \mathcal{H}} \pi_{h\tilde{h}} \hat{\psi}_{t+1,\tilde{h}} \right] + \beta \mathbb{E}_t \left[\Gamma_{t+1} (r_{t+1} - (1-\tau^K_{t+1}) \tilde{r}^K_{t+1}) \right] \\
& + \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[w_{t+1} \sum_{h \in \mathcal{H}} S_{t+1,h} (y_0^h)^\nu \left(\psi_{t+1,h} l_{t+1,h} + \lambda_{l,t+1,h} u'(c_{t+1,h}) \right) \right] \\
& + \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[\mu_{t+1} \left(Y_{t+1} \zeta_{t+1} - \frac{1}{\alpha} (r_{t+1} + \delta) K_t - w_{t+1} L_{t+1} \right) \right] \\
& - \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[\left((\gamma_{t+1} - \gamma_t) \Pi_{t+1} (\Pi_{t+1} - 1) - \frac{\varepsilon - 1}{\kappa} \gamma_{t+1} (\zeta_{t+1} - 1) \right) Y_{t+1} \right] \\
& - \beta \frac{\alpha - 1}{K_t} \mathbb{E}_t \left[(\tilde{r}^K_{t+1} + \delta) \left(\beta^{-1} \Upsilon_t + \Gamma_{t+1} (1 - \tau^K_{t+1}) K_t \right) \right] \\
& + \beta \frac{\alpha}{K_t} \mathbb{E}_t \left[Y_{t+1} \frac{1}{\sum_y S_y y^\nu} \left(1 - \zeta_{t+1} - \frac{\kappa}{2} (\Pi_{t+1} - 1)^2 \right) \sum_{h \in \mathcal{H}} S_{t+1,h} (y_0^h)^\nu \psi_{t+1,h} \right],
\end{aligned} \tag{118}$$

while for constrained agents, we have $a_{t,h} = -\bar{a}$ and $\lambda_{t,h} = 0$.

FOC with respect to $l_{t,h}$.

$$\begin{aligned}
\psi_{t,h} = {} & \frac{1}{w_t y_0^h} v'(l_{t,h}) + \frac{\lambda_{l,t,h}}{w_t y_0^h} v''(l_{t,h}) - (1-\alpha) \frac{\mu_t}{w_t L_t} (Y_t \zeta_t - w_t L_t) \\
& + (1-\alpha) \frac{Y_t}{w_t L_t} \left((\gamma_t - \gamma_{t-1}) \Pi_t (\Pi_t - 1) - \frac{\varepsilon - 1}{\kappa} \gamma_t (\zeta_t - 1) \right) \\
& + (1-\alpha) \frac{\tilde{r}^K_t + \delta}{w_t L_t} \left(\beta^{-1} \Upsilon_{t-1} + \Gamma_t (1 - \tau^K_t) K_{t-1} \right) \\
& + \alpha \frac{1}{L_t} \sum_{h \in \mathcal{H}} S_{t,h} y_0^h \left(l_{t,h} \psi_{t,h} + \lambda_{l,t,h} u'(c_{t,h}) \right) \\
& - (1-\alpha) \frac{Y_t}{w_t L_t} \frac{1}{\sum_y S_y y^\nu} \left(1 - \zeta_t - \frac{\kappa}{2} (\Pi_t - 1)^2 \right) \sum_{h \in \mathcal{H}} S_{t,h} (y_0^h)^\nu \psi_{t,h}.
\end{aligned}$$

The matrix representation to solve for the steady-state values of Lagrange multiplier, is similar to the one Section D.5, with one exception, which is the steady-state value of the Lagrange multiplier, γ, and which becomes:

$$\begin{aligned}
\frac{\varepsilon - 1}{\kappa} \gamma Y = {} & Y \left(\frac{(\boldsymbol{y}^\nu)^\top (\boldsymbol{S} \circ \boldsymbol{\psi})}{(\boldsymbol{y}^\nu)^\top \boldsymbol{S}} - \mu \right) + (\tilde{r}^K + \delta) \left(\beta^{-1} \Upsilon + \Gamma_t (1 - \tau^K) K \right) + w \mu L \\
& - w (\boldsymbol{y} \circ \boldsymbol{l})^\top (\boldsymbol{S} \circ \boldsymbol{\psi}) - w (\boldsymbol{y} \circ u'(\boldsymbol{c}))^\top (\boldsymbol{S} \circ \boldsymbol{\lambda}_l).
\end{aligned}$$

G.2 The quantitative exercise

We keep the same calibration as in Section 5, but with two modifications:

1. the inverse of IES is set to $\sigma = 1.4$, instead of 1.0 in the baseline calibration;

2. the Rotemberg cost parameter is set to $\kappa = 20$, instead of 100 in the baseline calibration. This implies a slope of the Phillips curve equal to 6%, as in Bhandari et al. (2021b).

Furthermore, the parameter ν driving how unequal the profit distribution is, is set to $\nu = 10$. This implies that the most productive agents hold 99.9% of profits.[27] The rest of the calibration is unchanged compared to Section 5 (including for the idiosyncratic and aggregate shocks).

These values are chosen so as to obtain a calibration closer to the one of Bhandari et al. (2021b) and ultimately to validate the findings of the sequence of simple models of Section 6. We recall that these findings were twofold: (i) in the absence of fiscal policy, the IES, the slope of the Phillips curve, and the unequal profit distribution were key to generate a sizable inflation response; (ii) a simple fiscal rule could prove to be sufficient to significantly reduce the inflation response, while being an efficient tool to provide insurance against the aggregate shocks.

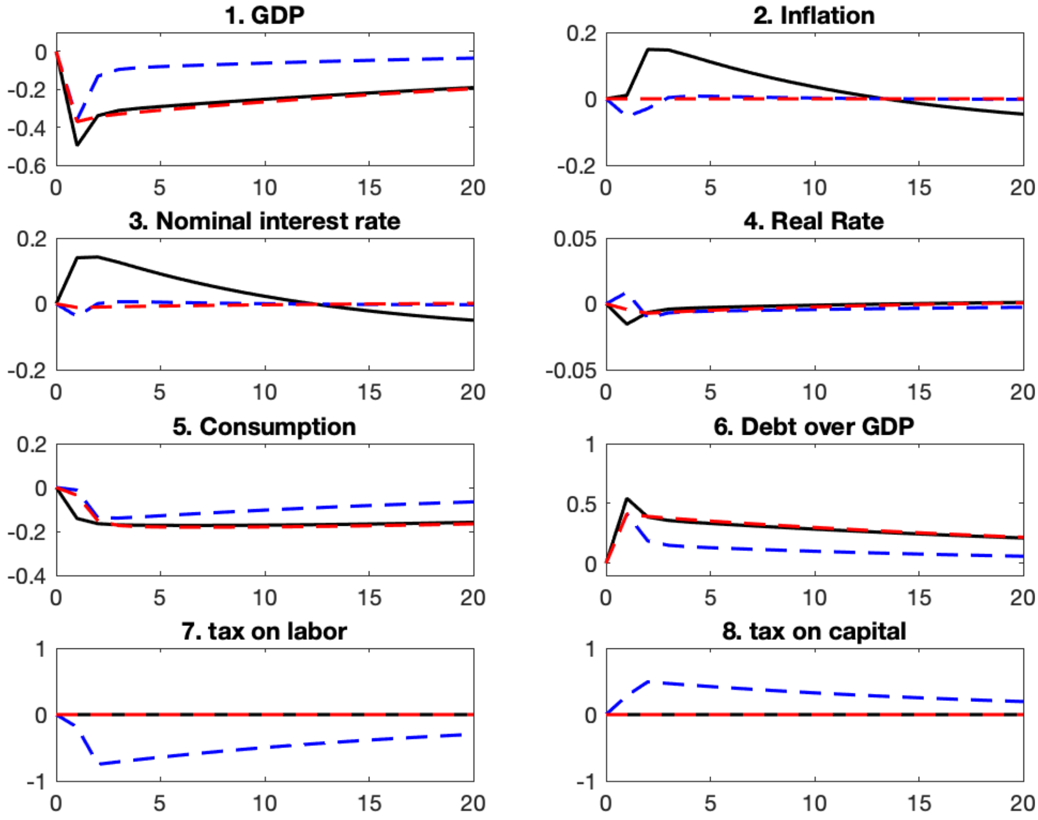


Figure 6: Impulse response function after a negative productivity shock for relevant variables. The black line is the benchmark economy with Fiscal Rule 1 and optimal monetary policy. The red dashed line is Economy 2, where we impose $\Pi_t = 1$. The blue dashed line is Economy 3 with time varying taxes and optimal monetary policy.

[27] We checked that increasing ν further has a negligible impact on model outcomes.

As in the main text, we consider three economies. Economies 1 and 2 are the same as in Section 5 and both implement Fiscal Rule 1, characterized by $(\sigma_1^L, \sigma_2^L, \sigma_1^K, \sigma_2^K, \sigma^T, \sigma^B) = (0, 0, 0, 0, 8.5, 4)$. Economy 1 considers the optimal inflation response (given fiscal policy 1), while Economy 2 sets a constant inflation $\Pi_t = 1$ for all dates. Economy 3 implements Fiscal Rule 3, corresponding to the set of parameters $(\sigma_1^L, \sigma_2^L, \sigma_1^K, \sigma_2^K, \sigma^T, \sigma^B) = (-0.6, -1.8, 0.9, 0.7, 8.5, 4)$, and where capital and labor tax rate are time-varying, and where inflation is optimal. The IRFs are plotted in Figure 6. This confirms the two findings of the simple models of Section 6.1. First, in the absence of time-varying taxes (but with Fiscal Rule 1), the maximal inflation response approximately amounts to 0.15%, which is close to the value of Bhandari et al. (2021b). Second, a simple fiscal rule can sizably decrease the inflation response, while offering better insurance, as can be seen through the aggregate consumption path, which overall dominates the no-rule case.

		Economy 1	Economy 2	Economy 3
Y	Mean	1.37	1.37	1.37
	Std(%)	1.41	1.42	0.47
C	Mean	0.94	0.94	0.94
	Std(%)	1.08	1.10	0.48
K	Mean	13.74	13.74	13.74
	Std(%)	1.61	1.63	0.19
L	Mean	0.38	0.38	0.38
	Std(%)	0.31	0.17	1.06
τ^L	Mean	0.28	0.28	0.28
	Std(%)	0.00	0.00	2.34
τ^K	Mean	0.36	0.36	0.36
	Std(%)	0.00	0.00	1.56
B	Mean	13.45	13.45	13.45
	Std(%)	0.16	0.16	0.19
T	Mean	0.11	0.11	0.11
	Std(%)	5.03	5.92	15.44
Π	Mean	1.00	1.00	1.00
	Std(%)	0.68	0.68	0.07
Correlations				
$corr(\Pi, Y)$		0.20	0.00	0.70
$corr(\tau^K, Y)$		0.00	0.00	-0.76
$corr(\tau^L, Y)$		0.00	0.00	0.69
$corr(B, Y)$		-0.96	-0.97	-0.81
$corr(C, Y)$		0.94	0.93	0.64
$corr(Y, Y_{-1})$		0.93	0.96	0.58
$corr(B, B_{-1})$		0.96	0.96	0.96

Table 9: First- and second-order moments for key variables, in the three economies (Economy 1 is the benchmark; Economy 2 is the economy with constant inflation rate; Economy 3 is the economy with time-varying fiscal policy).

These two points are also confirmed by the second-order moments of Table 9. It can be observed that Economy 3 – with a time-varying fiscal rule – features a lower volatility of GDP, aggregate consumption, and inflation, compared to Economy 1. The volatility reduction is sizable and is approximately twice smaller for aggregate variables and ten times smaller for inflation. The comparison of Economies 1 and 2 furthermore shows that despite a sizable inflation response, inflation is a rather inefficient insurance tool for redistribution, as the volatility reductions of GDP or aggregate consumption are modest – especially, compared to the fiscal rule.[28]

H Pareto weights and optimal fiscal policy

In this section, we explain how using Pareto weights in the social welfare function enables us to jointly consider optimal fiscal and monetary policies – with a missing instrument to avoid the equivalence result of Section 3. We assume here that the capital tax does not vary along the business cycle and is fixed at its optimal steady-state value τ^K_{SS}.

Instead of considering a utilitarian aggregate welfare, we consider a general welfare function that depends on the weights on the utility of each agent. We assume that these weights are consistent with the sequential representation, and depends on initial conditions and idiosyncratic and aggregate history. The weight of an agent $i \in \mathcal{I}$ is $\omega_{i,t} = \omega_t((a_{-1}, y_0), y_i^t, z^t)$, where the sequence of weights satisfy $1 = \int_i \omega_{i,t}\ell(di) = \sum_{y_i^t \in \mathcal{Y}^t} \sum_{y_0 \in \mathcal{Y}} \int_{a_{-1} \in [-\bar{a};+\infty)} \omega_t((a_{-1}, y_0), y_i^t, z^t)\theta_t(y_i^t)\Lambda(da_{-1}, y_0)$ for $t \geq 0$. These weights represent the relative importance of each agent in the planner's objective and will be calibrated in our quantitative exercise below, so as to match the US fiscal and monetary policies at the steady-state. Formally, the aggregate welfare criterion is:

$$\widetilde{W}_0 = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \omega_{i,t} \left(u(c_{i,t}) - v(l_{i,t}) \right) \ell(di) \right]. \tag{119}$$

We consider the same setup as in Section 4, except that taxes are not determined through rules (50)–(52), and obviously that the planner's objective is defined in (119). The Ramsey planner's program now involves the labor tax and can be written as:

$$\max_{\left(\tau_t^L, T_t, w_t, r_t, \tilde{w}_t, \tilde{r}_t^K, \tilde{R}_t^N, B_t, K_t, L_t, \Pi_t, (a_{i,t}, c_{i,t}, l_{i,t}, \nu_{i,t})_i\right)_{t \geq 0}} \widetilde{W}_0, \tag{120}$$

subject to: $\tau_t^k = \tau_{SS}^k$, and the same constraints as in Section 4.3 (but the fiscal rules).

To save some space, we directly solve for the truncated model.

[28]It is noteworthy to recall that the rule that we have chosen is not optimal by any criterion and a very large number of rules could provide a very similar conclusion.

H.1 The truncated Ramsey model

The truncated Ramsey program can be written as:

$$\max_{\left((a_{t,h},c_{t,h},l_{t,h})_{h\in\mathcal{H}},w_t,r_t,\tilde{w}_t,\tilde{r}_t,B_t,T_t,\Pi_t\right)_{t\geq 0}} \mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\sum_{h\in\mathcal{Y}}\Big[S_{t,h}\Big(\omega_h\left(\xi_h^{u,0}u(c_{t,h})-\xi_h^{v,0}v(l_{t,h})\right)$$
$$-(\lambda_{t,h}-\tilde{\lambda}_{t,h}(1+r_t))\xi_h u'(c_{t,h})-\lambda_{l,t,h}\left(\xi_{t,h}^{v,1}v'(l_{t,h})-w_t y_{t,h}\xi_h^{u,1}u'(c_{t,h})\right)\Big)\Big],$$

subject to truncated-history constraints:

$$c_{t,h}+a_{t,h}=w_t y_0^N l_{t,h}+(1+r_t)\tilde{a}_{t,h}+T_t,$$
$$\tilde{a}_{t,h}=\sum_{\tilde{h}\in\mathcal{H}}\pi_{\tilde{h}h}\frac{S_{t-1,\tilde{h}}}{S_{t,h}}a_{t-1,\tilde{h}},$$

to aggregate constraints:

$$G_t+r_t\left(B_{t-1}+K_{t-1}\right)+w_tL_t=B_t-B_{t-1}-T_t+\left(1-\kappa\frac{\pi_{2,t}}{2}\right)K_{t-1}^{\alpha}L_t^{1-\alpha}-\delta K_{t-1},$$
$$\Pi_t(\Pi_t-1)=\frac{\varepsilon-1}{\kappa}(\zeta_t-1)+\beta\mathbb{E}_t\Pi_{t+1}(\Pi_{t+1}-1)\frac{Y_{t+1}}{Y_t},$$
$$\zeta_t=\frac{1}{Z_t}\left(\frac{\tilde{r}_t^K+\delta}{\alpha}\right)^{\alpha}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{1-\alpha},\ A_t=K_t+B_t=\sum_{h\in\mathcal{Y}}S_{t,h}a_{t,h},\ L_t=\sum_{h\in\mathcal{Y}}S_{t,h}y_0^N l_{t,h},$$

and to interest rate constraints:

$$\mathbb{E}_t\left[\frac{R_t^N}{\Pi_{t+1}}\right]=\mathbb{E}_t\left[1+\tilde{r}_{t+1}^K\right],\ (r_t-(1-\tau_{SS}^K)\tilde{r}_t^K)A_{t-1}=(1-\tau_{SS}^K)(\frac{\tilde{R}_{t-1}^N}{\Pi_t}-1-\tilde{r}_t^K)B_{t-1}. \quad (121)$$

The Lagrangian associated to the previous Ramsey program is:

$$\mathcal{L}=\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\left[\sum_{h\in\mathcal{Y}}S_{t,h}\Big(\omega_h\left(\xi_h^{u,0}u(c_{t,h})-\xi_h^{v,0}v(l_{t,h})\right)\right.$$
$$-\sum_{h\in\mathcal{Y}}S_{t,h}(\lambda_{t,h}-\tilde{\lambda}_{t,h}(1+r_t))\xi_h u'(c_{t,h})-\sum_{h\in\mathcal{Y}}\lambda_{l,t,h}S_{t,h}\left(\xi_{t,h}^{v,1}v'(l_{t,h})-w_t y_{t,h}\xi_h^{u,1}u'(c_{t,h})\right)$$
$$-(\gamma_t-\gamma_{t-1})\Pi_t(\Pi_t-1)Y_t+\frac{\varepsilon-1}{\kappa}\gamma_t(\frac{1}{Z_t}\left(\frac{\tilde{r}_t^K+\delta}{\alpha}\right)^{\alpha}\left(\frac{\tilde{w}_t}{1-\alpha}\right)^{1-\alpha}-1)Y_t$$
$$+\mu_t\left(B_t-(1-\delta)B_{t-1}-G_t-T_t+(1-\frac{\kappa}{2}(\Pi_t-1)^2)Y_t-(r_t+\delta)A_{t-1}-w_tL_t\right)$$
$$+\frac{1}{\beta}\Upsilon_{t-1}(\frac{\tilde{R}_{t-1}^N}{\Pi_t}-(1+\tilde{r}_t^K))+\Gamma_t(r_tA_{t-1}-(1-\tau_t^K)(\tilde{r}_t^K K_{t-1}+(\frac{\tilde{R}_{t-1}^N}{\Pi_t}-1)B_{t-1})\Big)$$
$$\left.+\Lambda_t\left(\frac{\tilde{r}_t^K+\delta}{\alpha}K_{t-1}-\frac{\tilde{w}_t}{1-\alpha}L_t\right)\right].$$

H.2 First-order conditions

We define the net social value of liquidity, $\hat{\psi}$, as follows:

$$\hat{\psi}_{t,h} = \omega_h \xi_h^{u,0} u'(c_{t,h}) - \left(\lambda_{c,t,h}\xi_h^{u,E} - (1+r_t)\tilde{\lambda}_{c,t,h}\xi_h^{u,E} - \lambda_{l,t,h} w_t y_0^N \xi_h^{u,1}\right) u''(c_{t,h}) - \mu_t.$$

Derivative with respect to $\tilde{R}_t^N$.

$$(1-\tau_{SS}^K)\mathbb{E}_t\left[\frac{\Gamma_{t+1}}{\Pi_{t+1}}\right] B_t = \Upsilon_t \mathbb{E}_t\left[\frac{1}{\Pi_{t+1}}\right].$$

Derivative with respect to $\tilde{r}_t^K$.

$$\frac{1}{\beta}\Upsilon_{t-1} + \Gamma_t(1-\tau_{SS}^K)(A_{t-1} - B_{t-1}) = \frac{\varepsilon-1}{\kappa}\gamma_t K_{t-1} + \Lambda_t \frac{1}{\alpha} K_{t-1}.$$

Derivative with respect to Π_t.

$$0 = \mu_t \kappa(\Pi_t - 1) + (\gamma_t - \gamma_{t-1})(2\Pi_t - 1) - (\Gamma_t(1-\tau_{SS}^K)B_{t-1} - \beta^{-1}\Upsilon_{t-1})\frac{\tilde{R}_{t-1}^N}{Y_t \Pi_t^2}.$$

Derivative with respect to r_t.

$$\sum_{h\in\mathcal{H}} S_{t,h}\hat{\psi}_{t,h}\tilde{a}_{t,h} = -\sum_{h\in\mathcal{H}} S_{t,h}\tilde{\lambda}_{t,h}\xi_h U_c(c_{t,h}, l_{t,h}) - \Gamma_t A_{t-1}.$$

Derivative with respect to $\tilde{w}_t$.

$$\Lambda_t = (1-\alpha)\,\frac{\varepsilon-1}{\kappa}\gamma_t.$$

Derivative with respect to w_t.

$$0 = \sum_{h\in\mathcal{H}} S_{t,h} y_0^N (\hat{\psi}_{t,h} l_{t,h} + \lambda_{t,l,h}\xi_h^{u,1} u'(c_{t,h})).$$

Derivative with respect to B_t.

$$\begin{aligned}
\mu_t = {} & \beta\mathbb{E}_t\left[\mu_{t+1}\left(1-\delta+\alpha\frac{Y_{t+1}}{K_t}\left(1-\frac{\kappa}{2}(\Pi_{t+1}-1)^2\right)\right)\right] \\
& - \alpha\beta\mathbb{E}_t\left[((\gamma_{t+1}-\gamma_t)\Pi_{t+1}(\Pi_{t+1}-1) + \frac{\varepsilon-1}{\kappa}\gamma_{t+1})\frac{Y_{t+1}}{K_t}\right] \\
& + \frac{\varepsilon-1}{\kappa}\beta\mathbb{E}_t\left[\gamma_{t+1}\frac{\zeta_{t+1}Y_{t+1}}{K_t}\right] + \beta\mathbb{E}_t\Lambda_{t+1}\frac{\tilde{r}_{t+1}^K+\delta}{\alpha} \\
& + \beta(1-\tau_{SS}^K)\mathbb{E}_t\left[\Gamma_{t+1}\left(\frac{\tilde{R}_t^N}{\Pi_{t+1}} - 1 - \tilde{r}_{t+1}^K\right)\right].
\end{aligned}$$

Derivative with respect to $a_{t,h}$. For unconstrained truncated history h:

$$\psi_{t,h} = \beta\mathbb{E}_t\left[(1+r_{t+1})\psi_{i,t+1}\right] + \beta\mathbb{E}_t\left[\mu_{t+1}\left(-(1+r_{t+1}) + 1 - \delta + \alpha\frac{Y_{t+1}}{K_t}(1-\frac{\kappa}{2}(\Pi_{t+1}-1)^2)\right)\right]$$
$$+\beta\mathbb{E}_t\left[\Gamma_{t+1}\left(r_{t+1} - (1-\tau^K)\tilde{r}^K_{t+1}\right)\right] + \beta\mathbb{E}_t\left[\Lambda_{t+1}\frac{\tilde{r}^K_{t+1}+\delta}{\alpha}\right]$$
$$-\alpha\beta\mathbb{E}_t\left[\left((\gamma_{t+1}-\gamma_t)\Pi_{t+1}(\Pi_{t+1}-1) - \frac{\varepsilon-1}{\kappa}\gamma_{t+1}(\zeta_{t+1}-1)\right)\frac{Y_{t+1}}{K_t}\right]$$
$$+\beta\mathbb{E}_t\Lambda_{t+1}\left(\frac{\tilde{r}^K_{t+1}+\delta}{\alpha}\right),$$

while for constrained ones, we have $a_{t,h} = -\bar{a}$ and $\lambda_{t,h} = 0$. We take the difference with the FOC on B and find:

$$\hat{\psi}_{t,h} = \beta\left[(1+r_{t+1})\sum_{h\in\mathcal{Y}}\pi_{h\tilde{h}}\hat{\psi}_{t+1,h}\right] + \beta\mathbb{E}_t\left[\Gamma_{t+1}(r_{t+1} - (1-\tau^K_{t+1})(\frac{\tilde{R}^N_t}{\Pi_{t+1}} - 1))\right]. \quad (122)$$

FOC with respect to $l_{i,t}$.

$$\hat{\psi}_{t,h} = \frac{\omega_{t,h}\xi^v_{t,h}}{w_t y_h}v'(l_{t,h}) + \frac{\lambda_{l,t}\xi^v_{t,h}}{w_t y_h}v''(l_{t,h})$$
$$-(1-\alpha)\frac{\mu_t}{w_t L_t}Y_t\left(1-\frac{\kappa}{2}(\Pi_t-1)^2\right)$$
$$+(1-\alpha)\frac{1}{w_t L_t}\left((\gamma_t-\gamma_{t-1})\Pi_t(\Pi_t-1) - \frac{\varepsilon-1}{\kappa}\gamma_t(\zeta_t-1)\right)Y_t$$
$$-\Lambda_t\frac{\tilde{w}_t}{1-\alpha}\frac{1}{w_t}$$

FOC with respect to T_t.

$$\sum_{h\in\mathcal{H}} S_{t,h}\hat{\psi}_{t,h} \le 0, \quad (123)$$

with equality if $T_t > 0$.

H.3 Steady-state

We drop the t subscript to denote steady-state variables. The optimality conditions imply $\Pi = \zeta = 1$ and $0 = \gamma = \Gamma = \Upsilon = \Lambda$, while equations characterizing the Ramsey allocation are:

$$\hat{\psi}_h = \omega_h\xi^{u,0}_h u'(c_h) - \left(\lambda_{c,h}\xi^{u,E}_h - (1+r_t)\tilde{\lambda}_{c,h}\xi^{u,E}_h - \lambda_{l,h}w_t y^N_0\xi^{u,1}_h\right)u''(c_h) - \mu, \quad (124)$$

$$\hat{\psi}_h = \beta(1+r)\sum\pi_{h\tilde{h}}\hat{\psi}_{\tilde{h}},\ \text{if } h \text{ unconstrained},\ \lambda_h = 0,\ \text{otherwise}, \quad (125)$$

$$\hat{\psi}_h = \frac{1}{wy^N_0}(\omega_h\xi^{v,0}_h v'(l_h) + \lambda_{l,h}\xi^{v,1}_h v''(l_h)) - \mu(1-\alpha)\frac{Y}{wL}, \quad (126)$$

$$0 = \sum_{h\in\mathcal{Y}} S_h \left(\hat{\psi}_h \tilde{a}_h + \tilde{\lambda}_{c,h}\xi_h^{u,E} u'(c_h)\right) = \sum_{h\in\mathcal{Y}} S_h l_h y_h \left(\hat{\psi}_h + \lambda_{l,h}\xi_h^{u,1}(u'(c_h)/l_h)\right) = \sum_{h\in\mathcal{Y}} S_h\hat{\psi}_h, \quad (127)$$

$$1 = \beta(1+F_K). \quad (128)$$

H.4 A closed-form formula for the weights ωs

We use the same matrix notation as in Section D.5.

H.4.1 Notation

We start with the various ξs. Using truncation:

$$\xi_h^{u,0} = \frac{\sum_h S_h u(c_h)}{u(c_h)},\ \xi_h^{u,1} = \frac{\sum_h S_h u'(c_h)}{u'(c_h)},\ \xi_h^{v,0} = \frac{\sum_h S_h v(c_h)}{v(c_h)},$$

while as in Section D.5, we use Euler equation to determine $\boldsymbol{\xi}^{u,E}$ and $\boldsymbol{\xi}^{v,1}$ as:

$$\boldsymbol{\xi}^{u,E} = \left[\left(\boldsymbol{I} - \beta(1+r)\boldsymbol{\Pi}^\top\right)\boldsymbol{D}_{u'(\boldsymbol{c})}\right]^{-1}\boldsymbol{\nu}, \quad (129)$$

$$\boldsymbol{\xi}^{v,1} = w\boldsymbol{y}\circ\boldsymbol{l}\circ\boldsymbol{\xi}^{u,1}\circ u'(\boldsymbol{c})./v'(\boldsymbol{l}), \quad (130)$$

where ./ denotes the element-wise division for vectors. We also introduce the following notation:

$$\tilde{\boldsymbol{\xi}}^{v,1} := \boldsymbol{\xi}^{v,1}./(w\boldsymbol{y}),\ \tilde{\boldsymbol{\xi}}^{v,0} := \boldsymbol{\xi}^{v,0}./(w\boldsymbol{y}),\ \tilde{\boldsymbol{\xi}}^{u,1} := \boldsymbol{\xi}^{u,1}./\boldsymbol{l},$$

as well as: $\bar{\lambda}_l := \boldsymbol{S}\circ\lambda_l,\ \bar{\lambda}_c := \boldsymbol{S}\circ\lambda_c\ \bar{\psi} := \boldsymbol{S}\circ\hat{\psi},\ \bar{\Pi} := \boldsymbol{S}\circ\Pi^\top\circ(1./\boldsymbol{S}),\ \bar{\omega} := \boldsymbol{S}\circ\omega$. With this notation, we have: $\boldsymbol{S}\circ\tilde{\lambda}_c := \boldsymbol{\Pi}(\boldsymbol{S}\circ\boldsymbol{\lambda}_c) = \Pi\bar{\boldsymbol{\lambda}}_c$. These definitions imply that (124)–(128) become:

$$\bar{\boldsymbol{\psi}} = \bar{\boldsymbol{\omega}}\circ\boldsymbol{\xi}^{u,0}\circ u'(\boldsymbol{c}) - \mu\boldsymbol{S} \quad (131)$$
$$- \left(\bar{\boldsymbol{\lambda}}_c\circ\boldsymbol{\xi}^{u,E} - (1+r)\boldsymbol{\Pi}\bar{\boldsymbol{\lambda}}_c\circ\boldsymbol{\xi}^{u,E} - w\bar{\boldsymbol{\lambda}}_l\circ\boldsymbol{y}\circ\boldsymbol{l}\circ\tilde{\boldsymbol{\xi}}^{u,1}\right)\circ u''(\boldsymbol{c}),$$

$$\boldsymbol{P}\bar{\boldsymbol{\psi}} = \beta(1+r)\boldsymbol{P}\bar{\boldsymbol{\Pi}}\bar{\boldsymbol{\psi}}, \quad (132)$$

$$(\boldsymbol{I}-\boldsymbol{P})\bar{\boldsymbol{\lambda}}_c = 0, \quad (133)$$

$$\bar{\boldsymbol{\psi}} = \bar{\boldsymbol{\omega}}\circ\tilde{\boldsymbol{\xi}}^{v,0}\circ v'(\boldsymbol{l}) + \bar{\boldsymbol{\lambda}}_l\circ\tilde{\boldsymbol{\xi}}^{v,1}\circ v''(\boldsymbol{l}) - \mu F_L\boldsymbol{S}/w, \quad (134)$$

$$\tilde{\boldsymbol{a}}^\top\bar{\boldsymbol{\psi}} = -\left(\boldsymbol{\xi}^{u,E}\circ u'(\boldsymbol{c})\right)^\top\Pi\bar{\boldsymbol{\lambda}}_c, \quad (135)$$

$$(\boldsymbol{y}\circ\boldsymbol{l})^\top\bar{\boldsymbol{\psi}} = -\left(\boldsymbol{y}\circ\boldsymbol{l}\circ\tilde{\boldsymbol{\xi}}^{u,1}\circ u'(\boldsymbol{c})\right)^\top\bar{\boldsymbol{\lambda}}_l, \quad (136)$$

$$\boldsymbol{1}^\top\bar{\boldsymbol{\psi}} = 0, \quad (137)$$

$$\beta(1+F_K) = 1, \quad (138)$$

which is a system in $\bar{\boldsymbol{\psi}}, \bar{\boldsymbol{\omega}}, \bar{\boldsymbol{\lambda}}_c, \bar{\boldsymbol{\lambda}}_l, \mu$ – but (138), which actually characterizes the allocation.

H.4.2 FOCs in matrix form

Equation (134) yields:

$$\bar{\boldsymbol{\lambda}}_l = \boldsymbol{M}_0\bar{\boldsymbol{\omega}} + \boldsymbol{M}_1\bar{\boldsymbol{\psi}} + \mu\boldsymbol{V}_0, \tag{139}$$

with

$$\begin{aligned}
\boldsymbol{M}_0 &= -\boldsymbol{M}_1\boldsymbol{D}_{\tilde{\boldsymbol{\xi}}^{v,0}\circ v'(\boldsymbol{l})},\\
\boldsymbol{M}_1 &= \boldsymbol{D}^{-1}_{\tilde{\boldsymbol{\xi}}^{v,1}\circ v''(\boldsymbol{l})},\\
\boldsymbol{V}_0 &= F_L\boldsymbol{M}_1\boldsymbol{S}./w.
\end{aligned}$$

Then, equation (131) implies:

$$\bar{\boldsymbol{\psi}} = \hat{\boldsymbol{M}}_0\bar{\boldsymbol{\omega}} + \hat{\boldsymbol{M}}_1\bar{\boldsymbol{\lambda}}_c + \hat{\boldsymbol{M}}_2\bar{\boldsymbol{\lambda}}_l - \mu\boldsymbol{S}, \tag{140}$$

with:

$$\begin{aligned}
\hat{\boldsymbol{M}}_0 &= \boldsymbol{D}_{\boldsymbol{\xi}^{u,0}\circ u'(\boldsymbol{c})},\\
\hat{\boldsymbol{M}}_1 &= -\boldsymbol{D}_{\boldsymbol{\xi}^{u,E}\circ u''(\boldsymbol{c})}(\boldsymbol{I} - (1+r)\boldsymbol{\Pi}),\\
\hat{\boldsymbol{M}}_2 &= w\boldsymbol{D}_{\boldsymbol{y}\circ\boldsymbol{l}\circ\tilde{\boldsymbol{\xi}}^{u,1}\circ u''(\boldsymbol{c})}.
\end{aligned}$$

So using (140) and (140), we obtain:

$$\bar{\boldsymbol{\psi}} = \boldsymbol{M}_3\bar{\boldsymbol{\omega}} + \boldsymbol{M}_4\bar{\boldsymbol{\lambda}}_c + \mu\boldsymbol{V}_1, \tag{141}$$

with:

$$\begin{aligned}
\boldsymbol{M}_2 &= \boldsymbol{I} - \hat{\boldsymbol{M}}_2\boldsymbol{M}_1,\ \boldsymbol{M}_4 = \boldsymbol{M}_2^{-1}\hat{\boldsymbol{M}}_1,\\
\boldsymbol{M}_3 &= \boldsymbol{M}_2^{-1}(\hat{\boldsymbol{M}}_0 + \hat{\boldsymbol{M}}_2\boldsymbol{M}_0),\\
\boldsymbol{V}_1 &= \boldsymbol{M}_2^{-1}(\hat{\boldsymbol{M}}_2\boldsymbol{V}_0 - \boldsymbol{S}).
\end{aligned}$$

Then, using (132), (133), and (141), we get:

$$\bar{\boldsymbol{\lambda}}_c = \boldsymbol{M}_5\bar{\boldsymbol{\omega}} + \mu\boldsymbol{V}_2, \tag{142}$$

with:

$$\begin{aligned}
\tilde{\boldsymbol{R}}_5 &= -((\boldsymbol{I} - \boldsymbol{P}) + \boldsymbol{P}(I - \beta(1+r)\bar{\boldsymbol{\Pi}})\boldsymbol{M}_4)^{-1}\boldsymbol{P}(I - \beta(1+r)\bar{\boldsymbol{\Pi}}),\\
\boldsymbol{M}_5 &= \tilde{\boldsymbol{R}}_5\boldsymbol{M}_3,\ \boldsymbol{V}_2 = \tilde{\boldsymbol{R}}_5\boldsymbol{V}_1.
\end{aligned}$$

We then use equation (135), which becomes with (141) and (142):

$$\mu = -\boldsymbol{L}_1 \bar{\omega},$$

with:

$$\begin{aligned} C_1 &= \tilde{\boldsymbol{a}}^\top (\boldsymbol{V}_1 + \boldsymbol{M}_4 \boldsymbol{V}_2) + (\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi} \boldsymbol{V}_2, \\ \boldsymbol{L}_1 &= \left(\tilde{\boldsymbol{a}}^\top (\boldsymbol{M}_3 + \boldsymbol{M}_4 \boldsymbol{M}_5) + (\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}))^\top \boldsymbol{\Pi} \boldsymbol{M}_5 \right) / C_1. \end{aligned}$$

We deduce that from (139), (141), and (142):

$$\begin{aligned} \bar{\boldsymbol{\lambda}}_l &= \hat{\boldsymbol{M}}_6 \bar{\boldsymbol{\omega}}, \\ \bar{\boldsymbol{\lambda}}_c &= (\boldsymbol{M}_5 - \boldsymbol{V}_2 \boldsymbol{L}_1) \bar{\boldsymbol{\omega}}, \\ \bar{\boldsymbol{\psi}} &= \boldsymbol{M}_6 \bar{\boldsymbol{\omega}}, \end{aligned}$$

with:

$$\begin{aligned} \boldsymbol{M}_6 &= \boldsymbol{M}_3 + \boldsymbol{M}_4 (\boldsymbol{M}_5 - \boldsymbol{V}_2 \boldsymbol{L}_1) - \boldsymbol{V}_1 \boldsymbol{L}_1, \\ \hat{\boldsymbol{M}}_6 &= \boldsymbol{M}_0 + \boldsymbol{M}_1 \boldsymbol{M}_6 - \boldsymbol{V}_0 \boldsymbol{L}_1. \end{aligned}$$

The constraints (136) and (137) become:

$$\begin{aligned} \boldsymbol{L}_3 \bar{\boldsymbol{\omega}} &= \boldsymbol{L}_4 \bar{\boldsymbol{\omega}} = 0, \qquad (143) \\ \text{with: } \boldsymbol{Las}_3 &= (\boldsymbol{y} \circ \boldsymbol{l})^\top \boldsymbol{M}_6 + \left(\boldsymbol{y} \circ \boldsymbol{l} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c}) \right)^\top \hat{\boldsymbol{M}}_6, \\ \boldsymbol{L}_4 &= \boldsymbol{1}^\top \boldsymbol{M}_6. \end{aligned}$$

H.4.3 Constructing the Pareto weights

We assume that there are K distinct Pareto weights, ω^s, where K is the number of productivity levels. Define $\boldsymbol{M}_7$ as the $N_{tot} \times K$ matrix of elements in $\{0, 1\}$. The element of row h in column y is 1 if the current productivity of history h is y. We thus have:

$$\bar{\boldsymbol{\omega}} = \boldsymbol{D_S} \boldsymbol{M}_7 \boldsymbol{\omega}^s.$$

We define the Pareto weights as the ones that are the closest to utilitarian weights, such that constraints (143) hold. Formally, they are given as a solution of the following minimization problem:

$$\begin{aligned} &\min_{\omega} \| \boldsymbol{\omega}^s - \boldsymbol{1}_K \|^2, \\ &\text{s.t. } \boldsymbol{L}_3 \boldsymbol{D_S} \boldsymbol{M}_7 \boldsymbol{\omega}^s = \boldsymbol{L}_4 \boldsymbol{D_S} \boldsymbol{M}_7 \boldsymbol{\omega}^s = 0. \end{aligned}$$

Denoting by $2\mu_3$ and $2\mu_4$ the Lagrange multipliers on the two constraints, the FOCs imply:

$$\boldsymbol{\omega}^s = \mathbf{1}_K + \sum_{k=3}^{4} \mu_k (\boldsymbol{L}_k \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7)^\top, \tag{144}$$

which once substituted in constraints (143) yield:

$$\begin{bmatrix} \mu_3 \\ \mu_4 \end{bmatrix} = \boldsymbol{M}_8^{-1} \boldsymbol{V}_8,$$

$$\text{with: } \boldsymbol{M}_8 = \begin{bmatrix} \boldsymbol{L}_3 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \\ \boldsymbol{L}_4 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \end{bmatrix} \begin{bmatrix} \boldsymbol{L}_3 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \\ \boldsymbol{L}_4 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \end{bmatrix}^\top,$$

$$\boldsymbol{V}_8 = - \begin{bmatrix} \boldsymbol{L}_3 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \mathbf{1}_K \\ \boldsymbol{L}_4 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \mathbf{1}_K \end{bmatrix}.$$

Finally, from (144), we deduce the following expression for the Pareto weights:

$$\boldsymbol{\omega}^s = \mathbf{1}_K + \boldsymbol{M}_8^{-1} \boldsymbol{V}_8 \begin{bmatrix} \boldsymbol{L}_3 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \\ \boldsymbol{L}_4 \boldsymbol{D}_{\boldsymbol{S}} \boldsymbol{M}_7 \end{bmatrix}^\top.$$

I The simple model

I.1 The model without capital

We assume that there are two productivity levels $y_2 > y_1 > 0$ with an equal mass $1/2$ of agents for each productivity level. The transition matrix between productivity levels is denoted $(\pi_{i,j})_{i,j=1,2}$. Agents supply one unit of labor, and efficient labor supply is normalized to 1 such that $\frac{y_1+y_2}{2} = 1$. Agents with the same current productivity level are assumed to belong to the same island where they pool beginning-of-period wealth. The after-pooling wealth at date t is denoted $\tilde{a}_{i,t}$ for agents on island i with productivity level i. In each island, a family head maximizes the intertemporal welfare of all family members (that are identical because of wealth pooling). Denoting by $a_{i,t}$ the end-of-period wealth, we deduce:

$$\tilde{a}_{1,t} = \pi_{11} a_{1,t} + \pi_{21} a_{2,t},$$
$$\tilde{a}_{2,t} = \pi_{12} a_{1,t} + \pi_{22} a_{2,t}.$$

Using a guess-and-verify strategy, we assume that type-1 agents are credit-constrained: $a_{1,t} = -\bar{a}$ at all dates. We will later provide the conditions for this to be the case. As assets are in zero net-supply, the end-of-period market clearing is: $a_{1,t} + a_{2,t} = 0$. We deduce:

$$a_{1,t} = -a_{2,t} = -\bar{a},$$
$$\tilde{a}_{1,t} = -\tilde{a}_{2,t} = -(\pi_{11} - \pi_{21})\bar{a}.$$

In the absence of capital ($\alpha = 0$) and because of the normalization of the labor supply, the production sector is $Y_t = Z_t$, and we deduce from Section 2.3 the following expression for profits:

$$\Omega_t = \left(1 - \frac{w_t}{Z_t} - \frac{\kappa}{2}(\Pi_t - 1)^2\right) Z_t L_t,$$

where we used $\zeta_t = w_t/Z_t$. The Phillips curve becomes:

$$\Pi_t(\Pi_t - 1) = \frac{\varepsilon - 1}{\kappa}(\frac{w_t}{Z_t} - 1) + \beta \mathbb{E}_t\Big[\Pi_{t+1}(\Pi_{t+1} - 1)\frac{Y_{t+1}}{Y_t}\Big].$$

The individual budget constraints on the two islands can be expressed as follows ($i = 1, 2$):

$$a_i + c_{i,t} = \frac{R^N_{t-1}}{\Pi_t}\tilde{a}_i + w y_i + \frac{y_i^\nu}{\Psi}\Omega_t,$$

with $\Psi = y_1^\nu + y_2^\nu$. As the family head in each island cares equally about their island members, the two Euler equations can be written as:

$$u'(c_{1,t}) > \beta \mathbb{E}_t \frac{R^N_t}{\Pi_{t+1}}\left(\pi_{11}u'(c_{1,t+1}) + \pi_{12}u'(c_{2,t+1})\right), \tag{145}$$
$$u'(c_{2,t}) = \beta \mathbb{E}_t \frac{R^N_t}{\Pi_{t+1}}\left(\pi_{21}u'(c_{1,t+1}) + \pi_{22}u'(c_{2,t+1})\right).$$

The first Euler equation has to be a strict inequality for credit constraints to bind for type-1 agents. Using a utilitarian welfare function, the Ramsey program can be written as:

$$\max_{\left(c_{1,t},c_{2,t},w_t,R^N_t,\Pi_t\right)_{t\geq 0}} \frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (u(c_{1,t}) + u(c_{2,t})),$$
$$c_{1,t} = \frac{R^N_{t-1}}{\Pi_t}\tilde{a}_1 - a_1 + w_t y_1 + \frac{y_1^\nu}{\Psi}\left(1 - \frac{w_t}{Z_t} - \frac{\kappa}{2}\left(\Pi_t - 1\right)^2\right) Z_t,$$
$$c_{2,t} = \frac{R^N_{t-1}}{\Pi_t}\tilde{a}_2 - a_2 + w_t y_2 + \frac{y_2^\nu}{\Psi}\left(1 - \frac{w_t}{Z_t} - \frac{\kappa}{2}\left(\Pi_t - 1\right)^2\right) Z_t,$$
$$u'(c_{2,t}) = \beta \mathbb{E}_t \frac{R^N_t}{\Pi_{t+1}}\left(\pi_{21}u'(c_{1,t+1}) + \pi_{22}u'(c_{2,t+1})\right),$$
$$\Pi_t(\Pi_t - 1)Z_t = \frac{\varepsilon - 1}{\kappa}(\frac{w_t}{Z_t} - 1)Z_t + \beta \mathbb{E}_t\Big[\Pi_{t+1}(\Pi_{t+1} - 1)Z_{t+1}\Big].$$

Denoting by $\lambda_{2,t}$ the Lagrange multiplier on the Euler equation of agents 2, and by γ_t the Lagrange multiplier on the Phillips curve, the Lagrangian associated to the Ramsey program is:

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Bigg[(u(c_{1,t}) + u(c_{2,t})) - (\gamma_t - \gamma_{t-1})\Pi_t(\Pi_t - 1)Z_t + \frac{\varepsilon - 1}{\kappa}\gamma_t(\frac{w_t}{Z_t} - 1)Z_t, \tag{146}$$
$$- \left(\lambda_{2,t} - \frac{R^N_{t-1}}{\Pi_t}\pi_{22}\lambda_{2,t-1}\right) u'(c_{2,t}) - \left(0 - \frac{R^N_{t-1}}{\Pi_t}\pi_{21}\lambda_{2,t-1}\right) u'(c_{1,t})\Bigg],$$

with $\lambda_{2,-1} = 0$. We denote by $\tilde{\lambda}_{1,t} = \pi_{21}\lambda_{2,t-1}$ and $\tilde{\lambda}_{2,t} = \pi_{22}\lambda_{2,t-1}$ the beginning-of-period Lagrange multipliers in each state.

The FOC with respect to Π_t is:

$$0 = \left(\psi_{1,t}\frac{y_1^\nu}{\Psi} + \psi_{2,t}\frac{y_2^\nu}{\Psi}\right)\kappa(\Pi_t - 1)Z_t + \frac{R_{t-1}^N}{\Pi_t^2}\left(\psi_{1,t}\tilde{a}_{1,t} + \psi_{2,t}\tilde{a}_{2,t} + \tilde{\lambda}_{1,t}u'(c_{1,t}) + \tilde{\lambda}_{2,t}u'(c_{2,t})\right)$$
$$+ (\gamma_t - \gamma_{t-1})(2\Pi_t - 1)Z_t. \tag{147}$$

The FOC with respect to R_t^N is:

$$0 = \beta\mathbb{E}_t\frac{1}{\Pi_{t+1}}\left(\psi_{1,t+1}\tilde{a}_1 + \psi_{2,t+1}\tilde{a}_2 + \tilde{\lambda}_{1,t+1}u'(c_{1,t+1}) + \tilde{\lambda}_{2,t+1}u'(c_{2,t+1})\right). \tag{148}$$

The FOC with respect to w_t is:

$$0 = \left(y_1 - \frac{y_1^\nu}{\Psi}\right)\psi_{1,t} + \left(y_2 - \frac{y_2^\nu}{\Psi}\right)\psi_{2,t} + \frac{\varepsilon - 1}{\kappa}\gamma_t.$$

We deduce from equations (147) and (148) that the steady-state allocation implies $\Pi = 1$. We can then compute the steady-state allocation and then deduce the steady-state values of Lagrange multipliers from previous equations. We also check that type-1 agents are indeed credit-constrained in equilibrium, i.e., that inequality (145) holds. The model after a TFP shock can then be simulated using a first-order perturbation of the model equations.

I.2 The model with capital and taxes

We consider the program with capital and taxes. The environment with capital but no tax is a special case where taxes are set to 0. The production function is now $Y_t = Z_t K_t^\alpha$, as labor is still normalized to $L_t = 1$. Again, we guess-and-verify that agents 2 hold the capital stock, while agents 1 are credit-constrained. The pooling mechanism on islands now concerns both nominal and real assets.

The government finances a lump-sum transfer T_t by raising a distorting capital tax, τ_t^K, which can be time-varying. The post-tax gross real return on nominal bonds is $1 + (1 - \tau_t^K)(\frac{R_{t-1}^N}{\Pi_t} - 1)$, while the post-tax gross real return on capital is $1 + (1 - \tau_t^K)\tilde{r}_t^K$, where, as in Section 2, $\tilde{r}_t^K$ stands for the before-tax real interest rate on capital. The governmental budget constraint is:

$$T_t = \tau_t^K\left(\tilde{r}_t^K K_{t-1} + \left(\frac{R_{t-1}^N}{\Pi_t} - 1\right)(\tilde{a}_1 + \tilde{a}_2)\right), \tag{149}$$

where $\tilde{a}_1 + \tilde{a}_2 = 0$, because nominal assets are in zero net-supply. The real wage is denoted $\tilde{w}_t$. The two budget constraints and the nominal and real Euler equations can be expressed as follows:

$$
\begin{aligned}
-\bar{a}+c_{1,t} &= (1+(1-\tau_t^K)(\frac{R_{t-1}^N}{\Pi_t}-1))\tilde{a}_1+(1+(1-\tau_t^K)\tilde{r}_t^K)\tilde{K}_{1,t}+\tilde{w}_t y_1+\frac{y_1^\nu}{\Psi}\Omega_t+T_t,\\
\bar{a}+K_t+c_{2,t} &= (1+(1-\tau_t^K)(\frac{R_{t-1}^N}{\Pi_t}-1))\tilde{a}_2+(1+(1-\tau_t^K)\tilde{r}_t^K)\tilde{K}_{2,t}+\tilde{w}_t y_2+\frac{y_2^\nu}{\Psi}\Omega+T_t,\\
u'(c_{2,t}) &= \beta\mathbb{E}_t(1+(1-\tau_{t+1}^K)(\frac{R_t^N}{\Pi_{t+1}}-1))\left(\pi_{21}u'(c_{1,t+1})+\pi_{22}u'(c_{2,t+1})\right),\\
u'(c_{2,t}) &= \beta\mathbb{E}_t(1+(1-\tau_{t+1}^K)\tilde{r}_{t+1}^K)\left(\pi_{21}u'(c_{1,t+1})+\pi_{22}u'(c_{2,t+1})\right),
\end{aligned}
$$

with the pooling expressions: $\tilde{K}_{1,t}=\pi_{21}K_{t-1}$ and $\tilde{K}_{2,t}=\pi_{22}K_{t-1}$. The Phillips curve is still given by (7). We denote by $\lambda_{k,2,t}$ the Lagrange multiplier on the Euler equation for capital, and by $\lambda_{b,2,t}$ the Lagrange multiplier on the Euler equation for the nominal asset. The beginning-of-period values of Lagrange multipliers on capital are $\tilde{\lambda}_{k,i,t}=\pi_{2,i}\lambda_{k,2,t}$, for $i=1,2$.

The Lagrange multiplier on the definition of ζ is Υ_t, and the one on the governmental budget constraint (149) is μ_t. The Lagrangian associated to the planner's program is now:

$$
\begin{aligned}
\mathcal{L} &= \mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\Big[u(c_{1,t})+u(c_{2,t})\\
&+\tilde{\lambda}_{k,1,t}(1+(1-\tau_t^K)\tilde{r}_t^K)u'(c_{1,t})+\tilde{\lambda}_{b,1,t}(1+(1-\tau_t^K)(\frac{R_{t-1}^N}{\Pi_t}-1))u'(c_{1,t})\\
&-(\lambda_{k,2,t}-\tilde{\lambda}_{k,2,t}(1+(1-\tau_t^K)\tilde{r}_t^K))u'(c_{2,t})-(\lambda_{b,2,t}-\tilde{\lambda}_{b,2,t}(1+(1-\tau_t^K)(\frac{R_{t-1}^N}{\Pi_t}-1)))u'(c_{2,t})\Big]\\
&-\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\left[(\gamma_t-\gamma_{t-1})\Pi_t(\Pi_t-1)-\frac{\varepsilon-1}{\kappa}\gamma_t(\zeta_t-1)\right]Z_tK_{t-1}^\alpha\\
&+\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\Upsilon_t\left[\zeta_t-\frac{1}{\alpha Z_t}(\tilde{r}_t^K+\delta)K_{t-1}^{1-\alpha}\right]+\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t\mu_t\left[\tau_t^K\tilde{r}_t^KK_{t-1}-T_t\right].
\end{aligned}
$$

The first-order conditions associated to the Ramsey program can now be easily derived.

Derivative with respect to $\tilde{R}_t^N$.

$$
0=\mathbb{E}_t\frac{1-\tau_{t+1}^K}{\Pi_{t+1}}\left[\left(\psi_{t+1}^1\tilde{a}_1+\psi_{t+1}^2\tilde{a}_2\right)+\tilde{\lambda}_{b,2,t+1}u'(c_{2,t+1})+\tilde{\lambda}_{b,1,t+1}u'(c_{1,t+1})\right].
$$

Derivative with respect to $\tilde{r}_t^K$.

$$
\begin{aligned}
0=&(\psi_{1,t}\tilde{K}_{1,t}+\psi_{2,t}\tilde{K}_{2,t}+\tilde{\lambda}_{1,t}^Ku'(c_{1,t})+\tilde{\lambda}_{2,t}^Ku'(c_{2,t}))(1-\tau_t^K)(r_t^K+\delta)\\
&+\left(\frac{\varepsilon-1}{\kappa}\gamma_t-\mu_t^P\right)\alpha\zeta_tZ_tK_{t-1}^\alpha-(1-\alpha)\Upsilon_t\zeta_t+(r_t^K+\delta)\mu_t\tau^KK_{t-1}.
\end{aligned}
$$

Derivative with respect to $\tilde{w}_t$.

$$0 = (y_1\psi_{1,t} + y_1\psi_{2,t})\tilde{w}_t + \left(\frac{\varepsilon - 1}{\kappa}\gamma_t - (\frac{y_1^\nu}{\Psi}\psi_{1,t} + \frac{y_2^\nu}{\Psi}\psi_{2,t})\right)(1-\alpha)\zeta_t Z_t K_{t-1}^\alpha + (1-\alpha)\zeta_t \Upsilon_t.$$

Derivative with respect to Π_t.

$$\begin{aligned} 0 = (1-\tau_t^K)\frac{R_{t-1}^N}{\Pi_t^2}\left(\psi_{1,t}\tilde{a}^1 + \psi_{2,t}\tilde{a}^2 + \tilde{\lambda}_{b,1,t}u'(c_{1,t}) + \tilde{\lambda}_{b,2,t}u'(c_{2,t})\right) \\ + (\gamma_t - \gamma_{t-1})(2\Pi_t - 1)Z_t K_{t-1}^\alpha + \left(\frac{y_1^\nu}{\Psi}\psi_{1,t} + \frac{y_2^\nu}{\Psi}\psi_{2,t}\right)\kappa(\Pi_t - 1)Z_t K_{t-1}^\alpha. \end{aligned}$$

Derivative with respect to T_t.

$$\mu_t = \psi_{1,t} + \psi_{2,t}.$$

Derivative with respect to K_t.

$$\begin{aligned} \psi_{2,t} &= \beta\mathbb{E}_t(1 + (1-\tau_{t+1}^K)r_{t+1}^K)(\pi_{22}\psi_{2,t+1} + \pi_{21}\psi_{1,t+1}) \\ &- \alpha\beta\mathbb{E}_t\left[(\gamma_{t+1} - \gamma_t)\Pi_{t+1}(\Pi_{t+1} - 1) - \frac{\varepsilon - 1}{\kappa}\gamma_{t+1}(\zeta_{t+1} - 1)\right] Z_{t+1}K_t^{\alpha-1} \\ &+ \alpha\beta\mathbb{E}_t\left(\frac{y_1^\nu}{\Psi}\psi_{1,t+1} + \frac{y_2^\nu}{\Psi}\psi_{2,t+1}\right)(1 - \zeta_{t+1} - \frac{\kappa}{2}(\Pi_{t+1} - 1)^2)Z_{t+1}K_t^{\alpha-1} \\ &- \beta\mathbb{E}_t\Upsilon_{t+1}\frac{1-\alpha}{\alpha Z_{t+1}}(\tilde{r}_{t+1}^K + \delta)K_t^{-\alpha} + \beta\mathbb{E}_t\mu_{t+1}\tau_{t+1}^K\tilde{r}_{t+1}^K. \end{aligned}$$

The steady state corresponds to $\Pi = 1$, $T = \tau^K = 0$, and $1 + r^K = R^N$. We can then compute the steady-state allocation, as well as the steady-state values of Lagrange multipliers. The model dynamics can then be simulated using perturbation methods.

J Monetary shocks and inequality

In this section, we solve the model to study the effect of a 1% contractionary monetary shock on the dynamics of inequality using a standard Taylor rule. As shown in Section 4, the dynamics of inflation are sensitive to the choice of the fiscal rule. Then, to abstract from the effect of fiscal policy, we assume that there is no public spending need ($G = 0$), hence zero tax. The rest of the calibration follows Section F. Firms' profits are distributed to households and we set $\nu = 10$ as in Section G. Because firms profits are null at steady-state, the steady state is the same as in Section F. We check that the dynamics of inequality are in line with the empirical evidence of Coibion et al. (2017). The nominal gross and pre-tax interest rate $\tilde{R}_t^N$ is set according to the following simple Taylor rule: $\frac{\tilde{R}_t^N}{\tilde{R}^N} = \left(\frac{\Pi_t}{\Pi}\right)^{\phi_\Pi}\zeta_t^{\text{Taylor}}$, where: (i) $\log(\zeta_t^{\text{Taylor}}) = \rho^{\text{Taylor}}\log(\zeta_{t-1}^{\text{Taylor}}) + \varepsilon_t^{\text{Taylor}}$ is a persistent monetary policy shock with $\rho^{\text{Taylor}} = 0.5$; (ii) $\tilde{R}^N$ and Π are steady-state values, and (iii) the response of the nominal interest rate to inflation is equal to $\phi_\Pi = 1.5$. These are standard values in this literature (see Galí, 2015).

We report in Figure 7 the IRFs for the innovation to the Taylor rule, the rate of inflation, and output. The responses of aggregate quantities to this contractionary monetary policy shock are consistent with empirical evidence and the rest of the literature (Kaplan et al., 2018 among others). The real interest rate falls, depressing consumption and investment and causing inflation and output to fall.

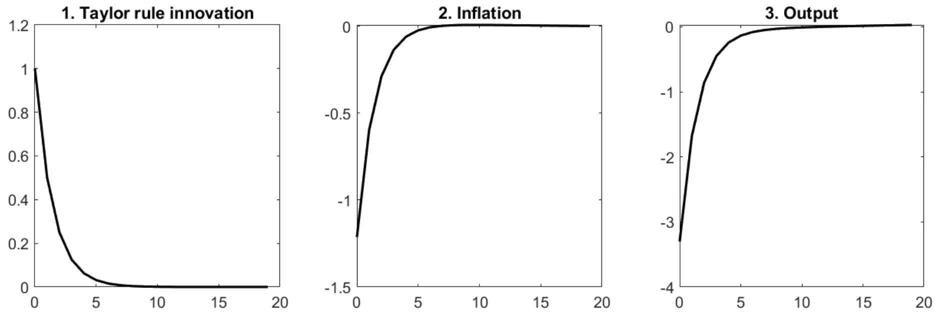


Figure 7: IRFs after a contractionary shock, in percentage deviation for all variables.

We also report in Figure 8 the responses of the Gini coefficients for income, consumption, and wealth corresponding to the same shock as in Figure 7. In line with the literature (see Gornemann et al., 2016 and Coibion et al., 2017), contractionary monetary policy increases these Gini coefficients and hence inequality.

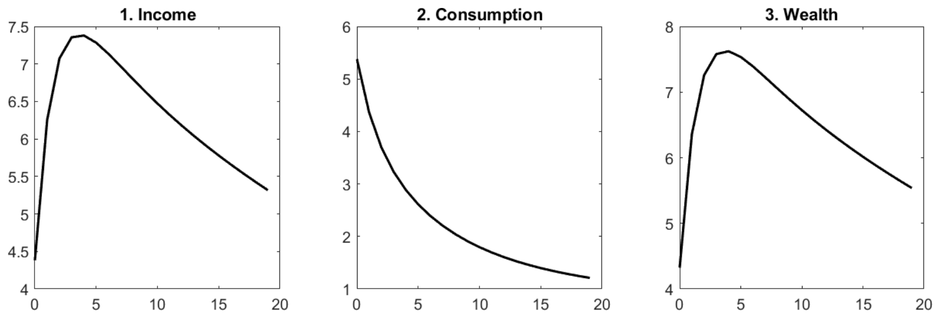


Figure 8: Gini coefficients of income, consumption, and wealth after the contractionary monetary shock of Figure 7. An increase by 1 indicates an increase of the coefficient by 0.1 pp.