Managing Inequality over Business Cycles: Optimal Policies with Heterogeneous Agents and Aggregate Shocks*

François Le Grand Xavier Ragot†

June 24, 2019

Abstract

We present a projection theory on the space of idiosyncratic histories for heterogeneous agent models. This method allows us to use a Lagrangian approach to solve for optimal Ramsey policies in heterogeneous agent models with aggregate shocks. It improves on current simulation methods by taking better advantage of steady-state information. We use this method to study the optimal level of unemployment insurance over the business cycle in a production economy. In the quantitative exercise, the optimal replacement rate is procyclical it and reduces aggregate consumption volatility.

Keywords: Incomplete markets, optimal policies, heterogeneous agent models.

JEL codes: E21, E44, D91, D31.

*A previous version of the paper circulated as "Optimal fiscal policy with heterogeneous agents and aggregate shocks". We thank Kurt Mitman for having shared his code with us. We also thank Edouard Challe, Varadarajan Chari, Wouter den Haan, Jonathan Heathcote, Christos Koulovatianos, Dirk Krueger, Felix Kubler, Etienne Lehmann, Albert Marcet, Alaïs Martin-Baillon, Ben Moll, Fabrizio Perri, Elisabeth Pröhl, Morten Ravn, Michael Reiter, Stephanie Schmitt-Grohe, Vincent Sterck, Martin Uribe, Gianluca Violante, Pierre-Olivier Weill, Sevin Yeltekin, and the seminar participants at the Federal Reserve Bank of Minneapolis, University of Zurich, University of Luxembourg, New York University, National Bank of Belgium, and European University Institute, at the "Macroeconomics of aggregate fluctuations and stabilization policy" conference, the T2M 2019, EEA 2018, SED 2019, and NBER SI 2017 MPF conference at the RED Conference in Barcelona. We thank Sumudu Kankamange and Thomas Weitzenblum for excellent research assistance. The research leading to these results received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement Integrated Macro-Financial Modeling for Robust Policy Design (MACFINROBODS, grant no. 612796).

†LeGrand: emlyon business school and ETH Zurich; legrand@em-lyon.com. Ragot: Sciences Po, OFCE and CNRS; xavier.ragot@gmail.com.

1 Introduction

Incomplete insurance market economies provide a useful framework for examining many relevant aspects of inequalities and individual risks. In these models, infinitely-lived agents face incomplete insurance markets and borrowing limits that prevent them from perfectly hedging their idiosyncratic risk, in line with the Bewley-Huggett-Aiyagari literature (Bewley 1983, Imrohoroğlu 1989, Huggett 1993, Aiyagari 1994, Krusell and Smith 1998). These frameworks are now widely used, since they fill a gap between micro- and macroeconomics, and enable the inclusion of aggregate shocks and a number of additional frictions on both the goods and labor markets. However, little is known about optimal policies in these environments due to the difficulties generated by the large and time-varying heterogeneity across agents. This is unfortunate, since a vast literature suggests that the interaction between wealth heterogeneity and capital accumulation has first-order implications for the design of optimal policies. In particular, the optimal design of time-varying unemployment benefits in an economy with fluctuating unemployment risk has not been studied in the general case yet, due to the difficulties generated by change in precautionary savings over the business cycle.

This paper presents a projection theory that can be used to derive optimal policies in incomplete insurance market economies with aggregate shocks. In incomplete insurance market economies, agents differ according to the full history of their idiosyncratic risk realizations. Huggett (1993) and Aiyagari (1994), using the results of Hopenhayn and Prescott (1992), have shown that economies without aggregate risk have a recursive structure when the distribution of wealth is introduced as a state variable. Unfortunately, the distribution of wealth has infinite support, which is at the root of many difficulties. Our main idea is to go back to the model sequential representation, so as to construct a projection theory on *any* finite partition of the set of idiosyncratic histories. This theory defines an exact economic model on a finite state space. Our model replicates the aggregate dynamics of the initial model, except that it is expressed using a finite number of "history bins" (groups of agents) instead of individual agents.

The theory is developed using two types of finite partitions. The first type, referred to as an *explicit partition*, involves a truncation in the space of idiosyncratic histories, such that agents with the same idiosyncratic history for the last N periods are assumed to belong to the same history bin. If there are k idiosyncratic states, the number of these N-histories – and thus the number of different bins – amounts to k^N. Interestingly, each history bin features time-varying, within-bin heterogeneity, as agents belonging to the same bin may have different histories $N+1$

periods ago. The projection on this first type of partition is simple to implement, but the number of bins grows exponentially with the truncation length. The second type of partition, referred to as an *implicit partition*, generalizes the previous construction to reduce the number of bins. The idea of implicit partitions involves constructing bins of individual histories from a partition of the steady-state wealth distribution through the one-to-one mapping between individual history and steady-state wealth (see Huggett 1993, for instance). Importantly, the implicit partition is a partition in the space of idiosyncratic histories. In the literature review below, we carefully explain the difference between the projection in the space of histories and other projection techniques, notably in the space of wealth used by Reiter (2009), for instance.

The key part of our projection theory is to properly aggregate individual choices (such as wealth, consumption, or savings) and individual constraints (such as Euler equations and budget constraints) at bin level. The resulting model is the so-called *projected model*, which is fully expressed in terms of the bins rather than the continuum of individual agents.[1] We show that the projected model replicates the aggregate dynamics of the model (in terms of prices and quantities, such as capital, labor, and output) thanks to the introduction of a finite number of time-varying *correcting coefficients* in Euler equations and budget constraints, which have intuitive interpretations.

However, the true model solution is needed to fully characterize the projected model. To make the projected model operational, we construct an *approximated model*, which is a projected model in which the correcting coefficients are fixed at their steady-state values. The approximated model thus assumes that the within-bin heterogeneity, although present, is not time-varying. We prove that the approximated model can be made arbitrarily close to the true model by choosing small enough bins. In addition, a small number of history bins provides a very good approximation of the dynamics of the model – whose outcomes cannot be distinguished from other solution techniques. Finally, in the case of explicit partitions, the approximated model can be micro-founded by a model with a specific form of insurance-market incompleteness, which may provide additional insights.

The advantage of this projection method is twofold. First, the approximated model has a simple structure with a finite number of bins. The tools developed in dynamic contracts, sometimes referred to as the Lagrangian approach and developed by Marcet and Marimon (2011), can thus be used to solve the Ramsey problem with an arbitrary set of instruments for

[1] In terms of wording, the *projection* of the model refers both to the reduction of the dimensionality of the state space by aggregation and the explicit derivation of laws of motion of the relevant aggregates. A bin is an element of the partition and is thus a set of idiosyncratic histories.

the planner. It is therefore possible to derive optimal policies with heterogeneous agents and aggregate shocks. The accuracy of the approximated model can easily be gauged, for instance, by studying the convergence properties of the correcting coefficients on increasingly finer partitions. The second advantage of the approximated model is that it is extremely fast. If we compare the different solution techniques, as in Den Haan (2010), a standard model with aggregate shocks can be solved in less than two seconds with a solver such as Dynare, while the outcomes do not significantly differ from those of the standard Krusell-Smith resolution.

We use projection theory to characterize the optimal unemployment benefits over the business cycle in the economy considered by Krueger, Mittman, and Perri (2018), which is a generalization of the economy studied in Krusell and Smith (1998). Agents face both productivity risk and time-varying employment risk. The economy is hit by aggregate shocks that affect technology and labor market transitions. Agents choose their labor supply when working, consume, save, and face incomplete markets for the idiosyncratic risk and credit constraints. In this economy, a planner chooses the level of unemployment benefits in each period, which must be fully financed by a distorting labor tax. Although the economic trade-off is the standard trade-off between insurance and incentives, this problem is very hard to solve in general equilibrium. The level of unemployment benefits directly affects agents' welfare as well as their saving decisions and the dynamics of interest rates and wages. The projection method enables this resolution. The main quantitative finding is that optimal unemployment benefits are time-varying and procyclical. The procyclicality of unemployment benefits corresponds to a procyclical distorting labor tax. Compared with an economy where the unemployment benefit is constant and fixed at its steady-state value, the time-varying replacement rate reduces output volatility by 13%.

Literature review. This paper first contributes to the literature on incomplete insurance market economies with aggregate shocks. There are now various types of techniques for solving these models. Our projection method is related to other perturbation and projection methods. The recent work of Boppart, Krusell, and Mitman (2018) shows that these perturbation methods are accurate approximations of the dynamics of such models in many relevant environments, compared with global solution techniques. In particular, the closest methods to ours are Reiter (2009). The main idea proposed in these two methods is to project the distribution of wealth on a finite set to simulate the model. The main difference compared with our model is that our solution involves projecting on the space of histories and not on the space of wealth. This allows us to construct an approximated model, which is crucial for solving Ramsey programs.

Another difference is that our method can be solved using standard computational tools such as Dynare, which offers some benefits in terms of implementation simplicity and computational speed. Finally, the partition in the space of idiosyncratic histories is different from partitions in the space of aggregate histories, which can be used as an approximation device (see for instance Chien, Cole, and Lustig (2011)). Indeed, our partitions is used to derive an exact representation of a projected model, which we simulate wih perturbation methods. To our knowledge, this paper is the first to use the partition in the space of idiosyncratic histories.

Second, our paper contributes to the recent literature on optimal Ramsay policies in heterogeneous agent models without aggregate shocks. An initial paper studying Ramsey allocation in a general setup is that of Aiyagari (1995), who provides a characterization of the optimal capital tax. Other papers, such as Aiyagari and McGrattan (1998) or Krueger and Ludwig (2016), derive optimal policies by maximizing the aggregate steady-state welfare, rather than by determining the optimal Ramsey policy. However, the steady-state welfare criterion does not account for transitions and we show that it can generate an allocation that significantly differs from the optimal Ramsey allocation. A significant step in the resolution of Ramsey programs in such set-ups is Açikgöz (2015), further developed in Açikgöz, Hagedorn, Holter, and Wang (2018), who use an explicit Lagrangian approach to derive the planner's first-order conditions. Dyrda and Pedroni (2016) compute optimal policies without considering the planner's first-order conditions, but instead by directly maximizing the intertemporal steady-state, which is computationally very intensive. Nuño and Moll (2018) consider a continuous-time framework in which they use the techniques of Ahn, Kaplan, Moll, Winberry, and Wolf (2017) to simplify the derivation of the planner's first-order conditions. The structure of our approximated model considerably simplifies the computation of optimal policies. As the state-space has a finite dimension, we show that steady-state Lagrange multipliers can be derived by simple matrix algebra, which is an additional advantage of our method.

Third, our paper contributes to the literature on optimal Ramsey policies in heterogeneous agent economies with aggregate shocks, which is currently rather thin. A first strategy used in the literature is to simplify the economy to analyze a tractable equilibrium. In these models the wealth distribution only has one or two mass points (McKay and Reis 2016, Bilbiie and Ragot 2017, and Challe 2018, among others). These models provide important economic insights but they cannot identify relevant properties related to the time-varying wealth distribution. Their quantitative relevance is thus hard to assess. To the best of our knowledge, the only paper deriving optimal Ramsey policy in a general environment with incomplete insurance markets and aggregate shocks is Bhandari, Evans, Golosov, and Sargent (2017). Their method relies on a "primal approach" in which credit constraints cannot be occasionally binding. They can be either always binding or never binding. The construction of our approximated model works well with occasionally binding credit constraints, which is the relevant case in many environments.

Finally, regarding the application, our paper contributes to the literature on optimal unemployment benefits. This literature is huge and a large part of it employs the sufficient-statistics approach (see the surveys of Chetty 2009, Chetty and Finkelstein 2013, and Kolsrud, Landais, Nilsson, and Spinnewijn 2018 for recent developments), based on partial-equilibrium analysis. A handful of papers introduce general equilibrium effects, such as Mitman and Rabinovich (2015), Landais, Michaillat, and Saez (2018) or Ábrahám, Brogueira de Sousa, Marimon, and Mayr (2019), but they focus on labor market externalities and not on saving distortions. To the best of our knowledge, the only paper analyzing optimal unemployment insurance in general equilibrium with saving choices is Krusell, Mukoyama, and Sahin (2010). To simplify the quantitative exercise, the authors perform a welfare analysis by comparing different steady-states with different levels of unemployment benefits. Instead, we solve for the time-varying solution of a general Ramsey problem in an economy with aggregate shocks.

The rest of the paper is organized as follows. In Section 2 we present the environment. In Section 3 we present the general projection theory in the space of idiosyncratic histories. In Section 4 we construct the approximated model and in Section 5 we derive optimal Ramsey policies and discuss the economic trade-off for optimal unemployment benefits over the business cycle. Section 6 sets out our quantitative analysis.

2 The economy

Time is discrete and indexed by $t = 0, 1, 2, ...$ The economy is populated by a continuum of agents of measure 1, distributed on an interval $\mathcal{I}$ according to a measure $\ell(\cdot)$. We follow Green (1994) and assume that the law of large numbers holds.

2.1 Preferences

In each period, agents derive utility from private consumption c and disutility from labor l. The period utility function, denoted by $U(c, l)$, is assumed to be of the Greenwood-Hercowitz-

Huffman (GHH) type, exhibiting no wealth effect for the labor supply:

$$U(c,l) = u\left(c - \chi^{-1}\frac{l^{1+1/\varphi}}{1+1/\varphi}\right), \tag{1}$$

where $\varphi > 0$ is the Frisch elasticity of labor supply, $\chi > 0$ scales labor disutility, and $u : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously derivable, increasing, and concave, with $u'(0) = \infty$. Our results do not rely on the GHH functional form and we could consider a more general utility function U. The algebra is however simplified, especially in the Ramsey program, because of the absence of a wealth effect for the labor supply.

Agents have standard additive intertemporal preferences, with a constant discount factor $\beta > 0$. They therefore rank consumption and labor streams, denoted respectively by $(c_t)_{t\geq 0}$ and $(l_t)_{t\geq 0}$, using the intertemporal utility criterion $\sum_{t=0}^{\infty} \beta^t U(c_t, l_t)$.

2.2 Risks

We consider a general setup where agents face an aggregate risk, a time-varying unemployment risk, and a productivity risk, as modeled by Krueger, Mittman, and Perri (2018). As will be clear in the quantitative analysis below, this general setup allows us to match the wealth distribution and the realistic dynamics of the labor market.[2]

Aggregate risk. The aggregate risk affects both aggregate productivity and unemployment risk. At a given date t, the aggregate state is denoted by z_t and takes values in the (possibly continuous) state space $\mathcal{Z} \subset \mathbb{R}^+$. We assume that the aggregate risk is a Markov process. The history of aggregate shocks up to time t is denoted by $z^t = \{z_0, \ldots, z_t\} \in \mathcal{Z}^{t+1}$. For the sake of clarity, for any random variable $X_t : \mathcal{Z}^{t+1} \to \mathbb{R}$, we will denote its realization in state z^t by X_t, instead of $X_t(z^t)$,

Employment risk. At the beginning of each period, each agent $i \in \mathcal{I}$ faces an uninsurable idiosyncratic employment risk, denoted by e_t^i at date t. The employment status e_t^i can take two values, e and u, corresponding to employment and unemployment, respectively. We denote by $\mathcal{E} = \{e, u\}$ the set of possible employment statuses. An employed agent with $e_t^i = e$ can freely choose her labor supply l_t^i. An unemployed agent with $e_t^i = u$ cannot work and will receive an unemployment benefit financed by a distorting tax on labor and will suffer from a fixed

[2] In contrast to Krueger, Mittman, and Perri (2018), we introduce endogenous labor supply, such that labor taxes are distorting. In addition, we simplify the economy, removing the age dimension and non-persistent productivity shocks, as they are not necessary for obtaining a realistic labor market and wealth outcomes.

disutility reflecting a domestic effort. These aspects are further described below. A history of idiosyncratic shocks up to date t for agent i is denoted by $e^{i,t} = \{e_0^i, \dots, e_t^i\} \in \mathcal{E}^{t+1}$.

The employment status $(e_t^i)_{t\geq 0}$ follows a discrete Markov process with transition matrix $M_t(z^t) \in [0,1]^{2\times 2}$, which is assumed to depend on the history of aggregate shocks up to date t. The job separation rate between periods $t-1$ and t is denoted by $\Pi_{eu}(z^t) = 1 - \Pi_{ee}(z^t)$, while $\Pi_{ue}(z^t) = 1 - \Pi_{uu}(z^t)$ is the job finding rate between $t-1$ and t. The time-varying transition matrix across employment statuses is therefore:

$$M_t(z^t) = \begin{bmatrix} \Pi_{uu}(z^t) & \Pi_{ue}(z^t) \\ \Pi_{eu}(z^t) & \Pi_{ee}(z^t) \end{bmatrix}. \tag{2}$$

We denote by $S_{u,t}$ and $S_{e,t}$ the implied population shares of unemployed and employed agents, respectively, with $S_{u,t} + S_{e,t} = 1$.

Productivity risk. Agents' individual productivity, denoted by y_t^i, is stochastic and takes values in a finite set $\mathcal{Y} \subset \mathbb{R}_+$. Large values in $\mathcal{Y}$ correspond to high productivities. The before-tax wage earned by an employed agent i is the product of the aggregate wage w_t, dependent on aggregate shock, of labor effort l_t^i, and of individual productivity y_t^i. The total before-tax wage is therefore $y_t^i w_t l_t^i$. An unemployed agent will also carry an idiosyncratic productivity level that will affect her unemployment benefits and her disutility level, denoted by ζ_y (for productivity $y \in \mathcal{Y}$), associated with domestic production.

The history up to date t of the productivity shocks of an agent i is denoted by $y^{i,t} = \{y_0^i, \dots, y_t^i\}$. The productivity status follows a first-order Markov process where the transition probability from state $y_{t-1}^i = y$ to $y_t^i = y'$ is constant and denoted by $\Pi_{yy'}$.[3] In particular, it is independent of the agent's employment status. We denote by S_y the share of agents endowed with individual productivity level y. This share is constant through time because of assumptions regarding transition probabilities $\Pi_{yy'}$.

The individual state of any agent i is characterized by her employment status and her productivity level. We will denote by $s_t^i = (e_t^i, y_t^i)$ the date-t individual status of any agent, whose possible values lie in the set $\mathcal{S} = \mathcal{E} \times \mathcal{Y}$. Finally, we denote by $s^{i,t} = \{s_0^i, \dots, s_t^i\}$ a history until period t. From transition probabilities for employment and productivity, one can derive the measure $\mu_t : \mathcal{S}^{t+1} \to [0,1]$, such that $\mu_t(s^t)$ is the measure of agents with history s^t in period t.

[3] We follow Krueger, Mittman, and Perri (2018) and denote all transitions by Π. They will only be distinguished by subscripts.

2.3 Production

The good is produced by a unique profit-maximizing representative firm. This firm is endowed with production technology that transforms, at date t, labor L_t (in efficient units) and capital K_{t-1} into Y_t output units of the single good. The production function F is a Cobb-Douglas function with parameter $\alpha \in (0,1)$ featuring constant returns-to-scale. The capital must be installed one period before production and the total productivity factor Z_t is stochastic. Denoting by $\delta > 0$ the constant capital depreciation, the net output Y_t is formally defined as follows:

$$Y_t = F(Z_t, K_{t-1}, L_t) = Z_t K_{t-1}^{\alpha} L_t^{1-\alpha} - \delta K_{t-1}, \tag{3}$$

where the total productivity factor is the exponential of the aggregate shock z_t: $Z_t = \exp(z_t)$.

The two factor prices at date t are the aggregate before-tax wage rate w_t and the capital return r_t. The profit maximization of the producing firm implies the following factor prices:

$$w_t = F_L(Z_t, K_{t-1}, L_t) \text{ and } r_t = F_K(Z_t, K_{t-1}, L_t). \tag{4}$$

2.4 Unemployment insurance

A benevolent government manages an unemployment insurance (UI) scheme, in which labor taxes are raised to finance unemployment benefits. As labor supply is endogenous, labor tax is distorting. The government thus faces the standard trade-off between insurance and efficiency.

At any date t, unemployed agents receive an unemployment benefit that is equal to a constant fraction of the wage the agent would earn if she were employed (with the same productivity level). The replacement rate being by denoted ϕ_t, the unemployment benefit of an agent i endowed with productivity y_t^i equals $\phi_t w_t y_t^i l_{t,e}^i$, where $l_{t,e}^i$ is the labor supply of a (fictive) employed agent with productivity y_t^i and wage rate w_t. We follow Krueger, Mittman, and Perri (2018) for this specification. From the agents' perspective, the replacement rate is an exogenous process that depends on the aggregate state $\phi_t = \phi_t\left(z^t\right)$.

Unemployment benefits are financed solely by the labor tax, which is paid by employed agents only. Taxes amount to a constant share τ_t of the employed agents' wage and this proportion is identical for all employed agents. The contribution τ_t is set such that the UI scheme budget is balanced at any date t, no social debt being allowed:

$$\phi_t w_t \int_{i \in \mathcal{U}_t} y_t^i l_{t,e}^i \ell(di) = \tau_t w_t \int_{i \in \mathcal{I} \setminus \mathcal{U}_t} y_t^i l_t^i \ell(di), \tag{5}$$

where $\mathcal{U}_t \subset \mathcal{I}$ is the set of unemployed agents at t and $\mathcal{I} \setminus \mathcal{U}_t$ the set of employed agents.

2.5 Agents' program and resource constraints

2.5.1 Sequential formulation

We consider an agent $i \in \mathcal{I}$. She can save in an asset that pays the gross interest rate $1+r_t$. She is prevented from borrowing too much and her savings must remain greater than an exogenous threshold denoted $-\bar{a}$. At date 0, the agent chooses the consumption $(c_t^i)_{t\geq 0}$, labor supply $(l_t^i)_{t\geq 0}$, and saving plans $(a_t^i)_{t\geq 0}$ that maximize her intertemporal utility, subject to a budget constraint and the previous borrowing limit. Formally, her program is, for a given a_{-1}^i:

$$\max_{\{c_t^i, l_t^i, a_t^i\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t^i - \chi^{-1}\frac{l_t^{i,1+1/\varphi}}{1+1/\varphi}\right) \tag{6}$$

$$c_t^i + a_t^i = (1+r_t)a_{t-1}^i + \Big((1-\tau_t)1_{e_t^i=e} + \phi_t 1_{e_t^i=u}\Big)\, l_t^i y_t^i w_t, \tag{7}$$

$$a_t^i \geq -\bar{a}, \tag{8}$$

where $1_{e_t^i=e}$ is an indicator function equal to 1 if the agent is currently employed ($e_t^i = e$) and to 0 otherwise. The budget constraint (7) is standard and the expression $((1-\tau_t)1_{e_t^i=e}+\phi_t 1_{e_t^i=u})l_t^i y_t^i w_t$ is a compact formulation for the net wage (i.e., after taxes and after unemployment benefits) of the agent i endowed with productivity y_t^i, depending on whether she is employed ($e_t^i = e$) or unemployed ($e_t^i = u$). Finally, initial wealth a_{-1}^i is given for any agent i.

We denote by $\beta^t \nu_t^i$ the Lagrange multiplier of the credit constraint of agent i. The Lagrange multiplier is obviously null when the agent is not credit-constrained. Taking advantage of the GHH utility function, the first-order conditions of an employed agent's program (6)–(8) can be written as:

$$u'(c_t^i - \chi^{-1}\frac{l_t^{i,1+1/\varphi}}{1+1/\varphi}) = \beta \mathbb{E}_t\left[(1+r_{t+1})u'(c_{t+1}^i - \chi^{-1}\frac{l_{t+1}^{i,1+1/\varphi}}{1+1/\varphi})\right] + \nu_t^i, \tag{9}$$

$$l_t^{i,1/\varphi} = \chi(1-\tau_t) w_t y_t^i 1_{e_t^i=e}, \tag{10}$$

and, for unemployed agents:

$$u'(c_t^i - \chi^{-1}\frac{\zeta_{y_t^i}^{1+1/\varphi}}{1+1/\varphi}) = \beta \mathbb{E}_t\left[(1+r_{t+1})u'(c_{t+1}^i - \chi^{-1}\frac{\zeta_{y_{t+1}^i}^{1+1/\varphi}}{1+1/\varphi})\right] + \nu_t^i. \tag{11}$$

The GHH utility function implies a very simple labor supply expression, which only depends on current productivity and the after-tax real wage. Unemployed agents supply no labor, but they earn unemployment benefits and suffer from disutility related to home production.

We now turn to economy-wide constraints. First, financial market clearing implies the following relationship:

$$\int_i a_t^i \ell(di) = K_t. \tag{12}$$

The clearing of the goods market implies that total consumption, comprising private individual consumption, private firm consumption, and public consumption, equals total supply, itself the sum of output and past capital:

$$\int_i c_t^i \ell(di) + K_t = Y_t + K_{t-1}. \tag{13}$$

Since every employed agent endogenously supplies labor, while unemployed agents do not work, the labor L_t in efficient units is defined as:

$$L_t = \int_{i \in \mathcal{I} \setminus \mathcal{U}_t} y_t^i l_t^i \ell\left(di\right). \tag{14}$$

Using the transition matrix M_t in equation (2), we deduce that the law of motion for the populations of employed and unemployed agents, denoted by $S_{e,t}$ and $S_{u,t}$ respectively, is:

$$S_{u,t} = 1 - S_{e,t} = \Pi_{eu,t} S_{e,t-1} + \Pi_{uu,t} S_{u,t-1}. \tag{15}$$

The constant share of agents S_y with productivity y verifies: $S_y = \sum_{y \in \mathcal{Y}} S_{y'} \Pi_{y'y}$. We can now formulate our equilibrium definition.

Definition 1 (Sequential equilibrium) *A sequential competitive equilibrium is a collection of individual allocations $\left(c_t^i, l_t^i, a_t^i\right)_{t \geq 0, i \in \mathcal{I}}$, of aggregate quantities $\left(K_t, L_t, Y_t\right)_{t \geq 0}$, of price processes $(w_t, r_t)_{t \geq 0}$, and of UI policy $(\tau_t, \phi_t)_{t \geq 0}$, such that, for an initial wealth distribution $\left(a_{-1}^i\right)_{i \in \mathcal{I}}$, and for initial values of capital stock $K_{-1} = \int_i a_{-1}^i \ell(di)$, and of the aggregate shock z_{-1}, we have:*

1. *given prices, individual strategies $\left(c_t^i, l_t^i, a_t^i\right)_{t \geq 0, i \in \mathcal{I}}$ solve the agent's optimization program in equations (6)–(8);*

2. *financial, labor, and goods markets clear at all dates: for any $t \geq 0$, equations (12), (13), and (14) hold;*

3. *the UI budget is balanced at all dates: equation (5) holds for all $t \geq 0$;*

4. *factor prices $(w_t, r_t)_{t \geq 0}$ are consistent with the firm's program (4).*

The goal of this paper is to determine the replacement rate process that generates the sequential equilibrium-maximizing aggregate welfare, using a utilitarian welfare criteria. This is a difficult question, as the replacement rate affects the saving decisions of all agents, the capital stock, and the price dynamics. We propose a solution that involves three steps. First, we provide the projection theory of the model in the space of idiosyncratic histories, using a finite partition of this space (Section 3), and then construct the projected model. Second, we construct a consistent approximated model, which has a finite state-space representation (Section 4). Finally, we show how to solve the Ramsey problem using the approximated model (Section 5).

3 The projection theory

3.1 Partitions

At any date t, each agent $i \in \mathcal{I}$ is uniquely characterized by her personal history of idiosyncratic risk realizations $s^{i,t} = (e^{i,t}, y^{i,t})$, which includes both employment and productivity risk histories. Obviously, the number of idiosyncratic risk histories grows exponentially over time and the steady-state is characterized by an unbounded number of idiosyncratic histories. The core idea of our projection method is to group agents in a finite number of bins, according to their idiosyncratic histories. Henceforth, the economy will be represented by this finite set of bins and not by the continuum of agents $\mathcal{I}$.

More formally, the set of bins is defined as a partition of the set of idiosyncratic histories. A partition $\mathcal{H}$ is a *finite* collection of sets of idiosyncratic histories, such that at any date t, an idiosyncratic history s^t belongs to one and exactly one element h of the partition $\mathcal{H}$: for any s^t, a unique $h \in \mathcal{H}$ exists, such that $s^t \in h$. An element $h \in \mathcal{H}$ will be called a *history bin* and can be thought of as a collection of individual histories – see our examples below for an explicit description. In the remainder of the paper, we will say that an agent belongs to $h \in \mathcal{H}$ at date t if her idiosyncratic history s^t belongs to h. Finally, we will say that a partition $\mathcal{H}$ is a refinement of a partition $\bar{\mathcal{H}}$ if any elements of $\bar{\mathcal{H}}$ are a union of elements of $\mathcal{H}$. In other words, $\mathcal{H}$ is finer than $\bar{\mathcal{H}}$.

To obtain this finite-state model representation, we *project* individual programs and first-order conditions onto the history partition $\mathcal{H}$, such that bin variables and the law of motions are consistent with each other. A preliminary, though crucial, remark is that each history bin $h \in \mathcal{H}$ embeds some time-varying heterogeneity among agents, since different idiosyncratic histories are represented with the same bin h. We derive below the formal way of accounting for this within-bin heterogeneity.

For the sake of concreteness, we now present two types of partitions, explicit partitions based on the truncation of idiosyncratic histories, and implicit partitions, based on the steady-state distribution of wealth.

3.1.1 Explicit partition

A first solution for constructing a partition consists in relying on the truncation of idiosyncratic histories of an exogenously given length $N > 0$: we regroup all agents who have the same idiosyncratic histories for the last N periods in the same bin. More precisely, each idiosyncratic history s^t is represented by the realizations of idiosyncratic status over the N consecutive previous periods, which is denoted by a vector $s^N = (s_{-N+1}, \ldots, s_0) \in \mathcal{S}^N$. A history bin h of the partition $\mathcal{H}$ is then the collection of all individual histories, whose realizations in the last N periods are identical and equal to an N-length vector $s^N \in \mathcal{S}^N$. Formally, the history bin h corresponding to the vector s^N can be written as $\bigcup_{t\geq 0}\{s^t \in \mathcal{S}^t : (s_{t-N+1}, \ldots, s_t) = s^N\}$. The bin h is therefore isomorphic to an N-length vector $s^N \in \mathcal{S}^N$ and the partition $\mathcal{H}$ itself is also isomorphic to the set $\mathcal{S}^N$ of idiosyncratic histories of length N. We can thus simply identify history bins h with N-length vectors $s^N \in \mathcal{S}^N$. We note $s^N \succeq \tilde{s}^N$, if s^N in period t is a possible continuation of $\tilde{s}^N$ in period $t-1$.[4] When an agent is in bin $\tilde{h}$, corresponding to the vector $\tilde{s}^N$ in periods $t-1$, the probability that she switches to bin h, corresponding to the vector s^N in period t, is denoted by $\Pi^S_{\tilde{s}^N s^N,t}$, with:

$$\Pi^S_{\tilde{s}^N s^N,t} = \Pi_{\tilde{y}_0 y_0}\Pi_{\tilde{e}_0 e_0,t}1_{s^N \succeq \tilde{s}^N}. \tag{16}$$

In this expression, the quantity $\Pi_{\tilde{y}_0 y_0}\Pi_{\tilde{e}_0,e_0,t}$ is the transition probability between idiosyncratic states $\tilde{s}_0 = (\tilde{y}_0, \tilde{e}_0)$ corresponding to bin $\tilde{h}$ and $s_0 = (y_h, e_h)$, corresponding to bin h.

Explicit partitions are intuitive partitions and are easy to implement, but they generate history bins with very heterogeneous sizes that depend on the relative persistence of individual states.[5] As a consequence, a relatively large N may be needed to properly capture the heterogeneity within certain history bins. The cost of this large N is a large number of bins –

[4] The continuation of a finite dimensional vector is associated with a truncation, and is thus different from the continuation of a whole history: $s^N \succeq \tilde{s}^N$ if $s_i^N = \tilde{s}_{i-1}^N$, $i = 1, ..., -N+1$.

[5] If we further add a pooling mechanism to explicit partitions, it is possible to formally derive an insurance-market structure to provide a micro-foundation for the approximated model. This uses a specific insurance-market structure to allow solely for insurance contracts on the truncated history. This construction is presented in the Technical Appendix, as we focus here on a more general representation of the projection method.

growing exponentially with N –, including some of very small size, whose contribution to the global dynamics is limited. For this reason, although the theory is simple in this case, we present another construction that relies on implicit partitions.

3.1.2 Implicit partitions

An alternative (and more abstract) definition of the partitions uses steady-state properties of state variables, such as the wealth distribution.[6] Implicit partitions generalize the concept of explicit partitions and enable us to follow a smaller number of bins. The basic idea of implicit partitions is to consider a partition of the steady-state wealth distribution and to gather in the same bin all of the agents whose idiosyncratic history yields a steady-state wealth belonging to the same wealth bin. Importantly, implicit partitions proceed in the space of histories and not in the space of wealth.

More formally, the construction of implicit partitions relies on the steady-state equilibrium, or equivalently on the equilibrium in the absence of aggregate shocks (i.e., $Z_t = 1$). Huggett (1993), building on Hopenhayn and Prescott (1992), has shown that this equilibrium is characterized by an invariant distribution, depending on the beginning-of-period wealth and the current idiosyncratic state. We use these two elements to define our implicit partition of idiosyncratic histories, based on a partition of the wealth space $[-\bar{a}, +\infty)$. The finite wealth partition, denoted by $\mathcal{B}$, is a finite collection of wealth bins $(b)_{b\in\mathcal{B}}$, such that:

$$[-\bar{a}, +\infty) = \cup_{b\in\mathcal{B}} b, \text{ and } b \cap b' = \emptyset \text{ for all } b \neq b'.$$

At any date t, the beginning-of-period wealth a^i_{t-1} of an agent i can be seen as a function of the idiosyncratic history up to date t, s^{t-1}. Formally, we have $a^i_{t-1} = a(s^{t-1})$, where $a : s^{t-1} \mapsto a(s^{t-1})$ defines a mapping between idiosyncratic histories and beginning-of-period wealth. Importantly, this mapping is well defined and one idiosyncratic history s^{t-1} corresponds to a unique beginning-of-period wealth $a(s^{t-1})$ and therefore to a unique wealth bin b, which is the sole element of $\mathcal{B}$, such that $a(s^{t-1}) \in b$.

The partition $\mathcal{H}$ of idiosyncratic histories will furthermore be parametrized by the current idiosyncratic state s and the wealth bin b. Loosely speaking, an idiosyncratic history s^t at date t will belong to a given history bin characterized by the idiosyncratic state s and the wealth bin b if: (i) the date-t idiosyncratic state is $s_t = s$ and (ii) s^{t-1}, the beginning-of-period wealth

[6] Other methods, such as Reiter (2009), use the steady-state distribution of wealth. We first derive the full theory and then compare the differences between the two methods in Section 4.3.

associated with the history up to date $t-1$, belongs to wealth bucket b: $a(s^{t-1}) \in b$. Formally, the wealth partition family is denoted by $(\mathcal{B}_s)_{s\in\mathcal{S}}$ and counts Card$\mathcal{S}$ elements. The partition $\mathcal{B}_s$ will correspond to histories whose current state is s. The partition $\mathcal{H}$ of idiosyncratic histories will be denoted as $\mathcal{H} = (h_{(s,b_s)})_{(s,b_s)\in\bigcup_{s'\in\mathcal{S}}\{(s',\mathcal{B}_{s'})\}}$. We use the subscript s in b_s to underline the dependence of the wealth partition in s. For any date $t \geq 0$, for any idiosyncratic state $s \in \mathcal{S}$, and any wealth bin $b_s \in \mathcal{B}_s$, we have:

$$s^t = (s^{t-1}, s_t) \in h_{(s,b_s)} \Leftrightarrow s_t = s \text{ and } a(s^{t-1}) \in b_s. \tag{17}$$

Since the idiosyncratic histories of a given bin $h_{(s,b_s)}$ are only defined through relationship (17), the partition $\mathcal{H}$ will be said to be *implicit*. As we will see later, relationship (17) is crucial for understanding that our construction is based on the partition of idiosyncratic histories – and not of wealth – but we do not actually need to invert it in our computational solution to identify actual idiosyncratic histories.

Note that two histories with two different current idiosyncratic states will belong to two different bins. As a consequence, we can unambiguously assign to any given history $h \in \mathcal{H}$, a unique current individual state $s_h = (y_h, e_h)$, where the current productivity level is denoted by y_h and the current employment status by e_h. Using this notation, we can express the transition probability from $\tilde{h}_{(\tilde{s},b_{\tilde{s}})}$ to $h_{(s,b_s)}$ at date t as follows:

$$\Pi^S_{\tilde{h}_{(\tilde{s},b_{\tilde{s}})}h_{(s,b_s)},t} = \Pi_{y_{\tilde{h}}y_h}\Pi_{e_{\tilde{h}}e_h,t}\widetilde{\Pi}^S_{\tilde{h}h,t}, \tag{18}$$

The quantity $\Pi_{y_{\tilde{h}}y_h}\Pi_{e_{\tilde{h}}e_h,t}$ corresponds to the transition probability at t of moving from the idiosyncratic state $\tilde{s} = (y_{\tilde{h}}, e_{\tilde{h}})$ to $s = (y_h, y_h)$. The quantity $\widetilde{\Pi}^S_{\tilde{h}h,t}$ corresponds to the fraction of agents switching from $\tilde{h}$ to h and reflects the heterogeneity in transition probabilities. It is equal to 1 in the absence of heterogeneity within bin $\tilde{h}$. But in general, bin $\tilde{h}$ can contain histories $s^t_1 \neq s^t_2$, which have different transition probabilities from bin h. Since bins are not explicitly known, the quantity $\widetilde{\Pi}^S_{\tilde{h}h,t}$ has no straightforward analytical expression – except in particular cases. We provide further insights on the probabilities $\Pi^S_{\tilde{h}_{(\tilde{s},b_{\tilde{s}})}h_{(s,b_s)},t}$ in the example below.

Comparing the transition probabilities for explicit partitions in (16) and for implicit partitions in (18) reveals that the expressions are very close to each other except for the term $\widetilde{\Pi}^S_{\tilde{h}h,t}$, which only appears for implicit partitions. An explicit partition can indeed be seen as a particular implicit partition in which, by construction, *all* histories of a given history bin h have the same transition probabilities. In other words, there is no heterogeneity in *transition probabilities* within history bins. In equation (18), the term $\widetilde{\Pi}^S_{\tilde{h}h,t}$ should therefore be equal to either 1 or 0 and we fall back to expression (16) for explicit partitions. This is, for instance, the case in the explicit partition example provided in Section 3.1.1.

Example of an implicit partition. Assume that agents face an employment risk only, such that there are only two possible idiosyncratic states: employed (e) or unemployed (u).[7] For the sake of simplicity, we assume throughout this example that job transition probabilities are constant. The partition $\mathcal{B}_e$ for currently employed agents is assumed to have only one element containing all histories: $\mathcal{B}_e = \{[-\underline{a}, \infty)\}$. The partition for unemployed agents contains three elements and is denoted by $\mathcal{B}_u = \{b_{1,u}, b_{2,u}, b_{3,u}\}$. In this simple example, we assume that when using these two wealth partitions and relationship (17), we obtain a partition of idiosyncratic histories with 4 bins, which we denote by $\{h_e, (h_{i,u})_{i=1,\dots,3}\}$ and which has the following characteristics. First, all employed agents end up in the same history bin h_e because of the particular partition $\mathcal{B}_e$. The element h_e will simply be denoted as $\{e\}$, where we use the same notation as for explicit partitions. Second, for unemployed agents, there are three history bins, which are denoted by $h_{1,u} = \{e, e, u\}$, $h_{2,u} = \{u, e, u\}$, and $h_{3,u} = \{u, u\}$ – again with the same notation. We further assume that only agents in history bin $h_{3,u} = \{uu\}$ – who were unemployed for the two last periods – are credit-constrained. We therefore obtain a partition $\mathcal{H} = \{h_e, h_{u,1}, h_{u,2}, h_{u,3}\}$ and a set of credit-constrained histories $\mathcal{C} = \{h_{u,3}\}$. Using idiosyncratic transition probabilities, we can compute transition probabilities $\left(\Pi^S_{hh'}\right)_{h,h' \in \mathcal{H}}$ between history bins. These probabilities are constant because job transition probabilities are constant. For instance, the transition probability between bins h_e and $h_{1,u}$ is a transition between states e and u for agents with the history $\{e, e\}$ for the two last periods. In other words, of the agents in history bin $h_e = \{e\}$, only those agents who were also employed before can transit to $h_{1,u}$, while those who were unemployed cannot transit to $h_{1,u}$ and can only possibly transit to $h_{2,u}$. Formally, the constant transition probability $\Pi_{h_e h_{1,u}}$ can be expressed as: $\Pi^S_{h_e h_{1,u}} = \Pi_{eu} \frac{S_{ee}}{S_e}$, where $\frac{S_{ee}}{S_e}$ is the (constant) share of agents having history ee among agents in h_e. This can further be simplified into $\Pi^S_{h_e h_{1,u}} = \Pi_{eu}\Pi_{ee}$ since $\frac{S_{ee}}{S_e} = \Pi_{ee}$ and $\widetilde{\Pi}_{h_e h_{1,u}} = \Pi_{ee}$ We can similarly express other transition probabilities.

In this simple example featuring implicit partitions, we can explicitly derive the within-bin transition probabilities. However, in actual models, these time-varying probabilities cannot be

[7] This projection theory can be seen as a generalization of previous work such as Challe, LeGrand, and Ragot (2013) or Challe, Matheron, Ragot, and Rubio-Ramirez (2017).

analytically characterized because the composition of history bins is not explicitly known. The construction of the approximated model below will deal with this difficulty.

3.2 Projection of the model

We now consider a partition $\mathcal{H}$ containing a finite set of history bins $h \in \mathcal{H}$. We explain how to project our economy onto this partition, which can be either implicit or explicit. Indeed, as explained before, explicit partitions can be seen as a particular case of implicit partitions. We first introduce general concepts of projection before applying the methodology to our model in Section 3.3.

3.2.1 Projecting variables: The basics

The first step of the projection consists in examining each variable of interest (such as consumption, asset holdings, labor supply, etc.) in order to determine its value for each history bin h in the partition $\mathcal{H}$. First, the size of a bin $h \in \mathcal{H}$ at date t corresponds to the measure of agents with an idiosyncratic history s^t belonging to bin h. Recall that the measure of idiosyncratic histories is denoted by μ_t. The population size, denoted by $S_{h,t}$, can thus formally be defined as: $S_{h,t} = \sum_{s^t \in h} \mu_t(s^t)$. As $\mathcal{H}$ is a partition of idiosyncratic histories, any agent in a bin $h \in \mathcal{H}$ in period t was in a bin $\tilde{h} \in \mathcal{H}$ in the previous period $t-1$. Since the transition probability between $\tilde{h}$ and h is denoted by $\Pi^S_{\tilde{h}h,t}$ and defined in equation (18), the evolution of history bin sizes follows a recursive definition that can be written as:[8]

$$S_{h,t} = \sum_{\tilde{h} \in \mathcal{H}} \Pi^S_{\tilde{h}h,t-1} S_{\tilde{h},t-1}. \tag{19}$$

Consider now the projection of a generic individual choice variable, which can be savings or consumption, for instance. In general, at date t, this variable can depend on the histories of idiosyncratic risk, s^t, and of aggregate risk, z^t. We will denote it by $X_t\left(s^t, z^t\right)$. The projection on the partition $\mathcal{H}$ consists in averaging the variable X_t over agents belonging to the same bin h. This value, denoted by $X_{h,t}(z^t)$ or simply $X_{h,t}$ for a bin $h \in \mathcal{H}$, is formally defined as:

$$X_{h,t} = \mathbb{A}_{h,t}[X_t] \equiv \frac{\sum_{s^t \in \mathcal{H}} X_t\left(s^t, z^t\right) \mu_t(s^t)}{S_{h,t}}. \tag{20}$$

We will denote by $\mathbb{A}_{h,t}[X_t]$ the projection of a variable X onto the history bin h at date t.

[8] A formal derivation of this projection formula – as well as all projection formulas used below – can be found in Section A of the Appendix.

3.2.2 Projecting variables: Other mechanisms

Equations (19) and (20) provide the basic mechanisms for projecting the model onto the partition $\mathcal{H}$. However, other operations are also useful, although slightly more subtle.

First, we examine the projection of a transformed variable. Since the projection operations rely on the linear operator $\mathbb{A}_{h,t}$, the projection of a transformed variable generally differs from the transformation of the projected variable. In other words, $\mathbb{A}_{h,t}[f(X_t)]$ generally differs from $f(\mathbb{A}_{h,t}[X_t])$, except in particular cases – such as f being affine or X_t having a degenerate distribution. Because of this non-linearity, we will define the projection as:

$$\mathbb{A}_{h,t}[f(X_t)] = \xi^f_{h,t} f(X_{h,t}), \tag{21}$$

where the quantity $\xi^f_{h,t}$ embeds the non-linearity correction of f and the heterogeneity in X within the bin h. This correction is time-varying (because the distribution within bin h is in general time-varying) and depends on the function f and on the variable X. The parameter $\xi^f_{h,t}$ will be a useful tool in our numerical algorithm.

Second, we are interested in the projection of a past variable X_{t-1} onto a history h at date t. This is, for instance, useful for the projection of the beginning-of-period wealth of agents in the budget constraint. This projection is the average of a variable X_{t-1} over agents belonging to bin h in period t, but possibly to another bin $\tilde{h}$ in period $t-1$. These transitions have to be taken into account in the projection. More subtly, the transition probabilities for X generally differ from the transition probabilities $(\Pi^S_{\tilde{h}h,t})$ between the history bins defined in equation (18). The reason is that two agents with respective histories $s^{1,t} \neq s^{2,t}$, belonging to the same bin $\tilde{h}$ at date t, can face different probabilities for transitioning to bin h and, at the same time, can be endowed with different values for X. The combination of these two sources of heterogeneity implies that averaging within bin $\tilde{h}$ for the flow of the variable X generates transition probabilities for X that can differ from those for agents' flows (Π^S in our notation). We provide an illustrative example below. The projection $\mathbb{A}_{h,t}[X_{t-1}]$ at date t on history bin h of the lagged variable X_{t-1} can be written as:

$$\mathbb{A}_{h,t}[X_{t-1}] = \frac{\sum_{\tilde{h}\in\mathcal{H}} \Pi^X_{\tilde{h}h,t} S_{\tilde{h},t-1} X_{\tilde{h},t-1}}{S_{h,t}}, \tag{22}$$

where $(\Pi^X_{\tilde{h}h,t})_{\tilde{h},h\in(\mathcal{H}\times\mathcal{H})}$ is a probability transition matrix, for which we have $\Pi^X_{\tilde{h}h,t} \in [0,1]$ and $\sum_{h\in\mathcal{H}} \Pi^X_{\tilde{h}h,t} = 1$. An exact expression of the matrix $(\Pi^X_{\tilde{h}h,t})$ can be found in Appendix A. In general, this matrix depends on the variable X under consideration. It is noteworthy that in the absence of heterogeneity either for X within bin $\tilde{h}$ or for agents' transition probabilities, the transition probability $\Pi^X_{\tilde{h}h,t}$ becomes equal to $\Pi^S_{\tilde{h}h,t}$ – the transition probability (18) of agents between bins. This is, for instance, the case for explicit partitions.

Example. This example continues the one presented in Section 3.1.2 for implicit partitions. We aim to compute the projection $\mathbb{A}_{h_{1,u},t}[a_{t-1}]$ of wealth a_{t-1} on the history bin $h_{1,u} = \{e, e, u\}$ at t. All agents in bin $h_{1,u}$ come from bin $h_e = \{e\}$, and the transition probability between both bins – which we already computed – is $\Pi_{h_e h_{1,u}} = \Pi_{eu}\Pi_{ee}$. For wealth projection, it is useful to observe that bin h_e can be partitioned into two sub-bins, $h_{ee} = \{ee\}$ and $h_{ue} = \{ue\}$, depending on the employment status in the previous period. Only agents from the sub-bin $h_{ee} = \{ee\} \subset h_e$, i.e., those who have been employed for two consecutive periods, will transition to $h_{1,u}$ with probability Π_{ee}, while agents of h_{ue} have a zero probability of transitioning to $h_{1,u}$. Using the definitions of conditional probabilities and of these two sub-bins, we therefore note that the projection $\mathbb{A}_{h_{1,u},t}[a_{t-1}]$ can be expressed as: $\mathbb{A}_{h_{1,u},t}[a_{t-1}] = \frac{1}{S_{h_{1,u},t}} \sum_{s^{t-1} \in h_{ee}} X_{t-1}\left(s^{t-1}\right) \Pi_{eu} \mu_{t-1}(s^{t-1})$. We can simplify this expression further using equation (20):

$$\mathbb{A}_{h_{1,u},t}[a_{t-1}] = \Pi_{eu} \frac{S_{ee}}{S_{h_{1,u}}} a_{ee,t-1} = \Pi_{ee}\Pi_{eu} \frac{S_e}{S_{h_{1,u}}} a_{ee,t-1}, \tag{23}$$

where the quantity $a_{ee,t-1}$ generally differs from $a_{h_e,t-1}$. We in fact have $a_{h_e,t-1} = \frac{S_{ee}}{S_e} a_{ee,t-1} + \frac{S_{ue}}{S_e} a_{ue,t-1}$, which coincides with $a_{ee,t-1}$ only in the special case where $a_{ee,t-1} = a_{ue,t-1}$, i.e., in the absence of within-bin heterogeneity. From equations (22) and (23), we therefore define $\Pi^a_{h_e h_{1,u},t-1} = \Pi_{ee}\Pi_{eu} \frac{a_{ee,t-1}}{a_{e,t-1}}$, which in general differs from $\Pi_{h_e h_{1,u}} (= \Pi_{ee}\Pi_{eu})$, unless there is no heterogeneity in bin h_e.

To conclude this section on projection mechanisms, we observe that projecting expectations also implies the introduction of corrective elements. We will describe this procedure when dealing with the Euler equation in Section 3.3.

3.3 The projected model

We consider the economic model presented in Section 2 and a finite-size partition $\mathcal{H}$. The model is characterized by the following set of equations: (i) individual budget constraints, (ii) Euler equations, (iii) market clearing conditions, and (iv) the balance of the unemployment scheme. We will project the model and obtain the counterparts of these different equations for the partition $\mathcal{H}$.

Budget constraints. The individual budget constraint (7) is relatively straightforward to project and proceeds from the basic projection mechanism (20) and from the projection (22) for a lagged variable. Using proper notation, the budget constraint for any bin $h \in \mathcal{H}$ can be expressed as:

$$c_{h,t} + a_{h,t} \leq (1 + r_t) \sum_{\tilde{h} \in \mathcal{H}} \Pi^a_{\tilde{h}h,t} \frac{S_{\tilde{h},t-1}}{S_{h,t}} a_{\tilde{h},t-1} + ((1 - \tau_t) 1_{e_h = e} + \phi_t 1_{e_h = u}) \, l_{h,t} y_h w_t. \tag{24}$$

The interpretation of the projected budget constraint (24) is immediate: resources, comprising labor income and saving payoffs, are used to consume and save. The only subtlety relates to the projection of past savings, as already discussed.

Market clearing conditions. Projecting market clearing conditions is also relatively straightforward, as these conditions are linear. Equations (12)–(14), which state the market clearing conditions for capital, goods, and labor respectively, become:

$$K_t = \sum_{h \in \mathcal{H}} S_{h,t} a_{h,t}, \tag{25}$$

$$\sum_{h \in \mathcal{H}} S_{h,t} c_{h,t} + K_t = F(K_{t-1}, L_t) + K_{t-1}, \tag{26}$$

$$L_t = \sum_{h \in \mathcal{H}} S_{h,t} y_h l_{h,t}. \tag{27}$$

Again, the interpretation of partition-clearing conditions (25)–(27) is relatively direct: aggregate variables are equal to the sum of bin variables, weighted by bin size.

Unemployment insurance. Using individual labor Euler conditions (10), the UI budget constraint (5) becomes: $\phi_t \int_{i \in \mathcal{U}_t} \left(y^i_t\right)^{1+\varphi} \ell(di) = \tau_t \int_{i \in \mathcal{I} \setminus \mathcal{U}_t} \left(y^i_t\right)^{1+\varphi} \ell(di)$. We observe that the budget balance only depends on the current idiosyncratic state. This can therefore be simplified into $\phi_t \sum_{y \in \mathcal{Y}} S_{u,t} S_y y^{1+\varphi} = \tau_t \sum_{y \in \mathcal{Y}} S_{e,t} S_y y^{1+\varphi}$, or:

$$\phi_t S_{u,t} = \tau_t S_{e,t}, \tag{28}$$

which is independent of the partition in this economy.

First-order conditions. The first-order condition (10) for labor is straightforward to project, as it is linear. It implies that the labor supply for a bin h can be expressed as:

$$l_{h,t} = \chi^\varphi (1 - \tau_t)^\varphi w^\varphi_t y^\varphi_h 1_{e_h = e}. \tag{29}$$

Only employed agents supply labor and their effort depends solely on the labor tax (negatively) and on bin productivity (positively).

The projection of the Euler equation for consumption is more involved for two reasons: (i) the non-linearity of marginal utilities and (ii) the conditional expectation operator. The consumption Euler equation for agents can be expressed as:

$$\xi^u_{h,t} U_c(c_{h,t}, l_{h,t}) - \nu_{h,t} = \beta \mathbb{E}_t \left[(1 + r_{t+1}) \sum_{\tilde{h} \in \mathcal{H}} \Pi^u_{h\tilde{h},t+1} \xi^u_{\tilde{h},t+1} U_c(c_{\tilde{h},t+1}, l_{\tilde{h},,t+1}) \right], \tag{30}$$

where $\nu_{h,t} = 0$ if agents in h are not credit-constrained. First, the quantities $\xi^u_{h,t}$ in (30) correct for the nonlinearity of U_c and are defined in (21). Second, the expectation operator $\mathbb{E}_t[\cdot]$ should be understood with respect to aggregate risk only, since individual risks are handled explicitly in a developed summation. The terms $(\Pi^u_{h\tilde{h},t})$ are nonnegative and they aggregate future transitions across history bins. As for the projection of lagged variables, the term $\Pi^u_{h\tilde{h},t+1}$ reduces to $\Pi^u_{h\tilde{h},t+1} = \Pi^S_{h\tilde{h},t+1}$ in the absence of bin heterogeneity, which greatly simplifies the projected Euler equation (30).

4 The approximated model

The projected model in Section 3 is an exact representation of the initial model, except that it follows history bins rather than individual agents. Time-varying, within-bin heterogeneity is captured by correcting parameters, such as $\Pi^a_{h\tilde{h},t}$, $\Pi^u_{h\tilde{h},t}$, and $\xi^u_{h,t}$. These correcting coefficients are further characterized in the approximated model below.

To track the dynamics of these coefficients in the presence of aggregate shocks, we need to solve the full model to follow the time-varying heterogeneity within each bin. The approximated model is based on the assumption that the value of these parameters is not time-varying and is equal to their steady-state values. As a result, although we consider heterogeneity in each bin, this heterogeneity is not considered to be time-varying. We thus solve for the dynamics of an approximate model where the correcting parameters $(\Pi^a_{h\tilde{h},t}, \Pi^u_{h\tilde{h},t}, \xi^u_{h,t})$ are introduced to match both steady-state transitions and distributions.

4.1 Steady-state economy

The equations characterizing the choices at the bin level are: the Euler equations for consumption and labor (29) and (30), the budget constraint (24), and the dynamics of bin sizes (19). At the steady-state, dropping the subscript t, these equations become respectively, for all $h \in \mathcal{H}$:

$$\xi_h^u U_c\left(c_h, l_h\right) - \nu_h = \beta(1+r) \sum_{\tilde{h}\in\mathcal{H}} \Pi_{h\tilde{h}}^u \xi_{\tilde{h}}^u U_c(c_{\tilde{h}}, l_{\tilde{h}}), \tag{31}$$

$$l_h = \chi^{\varphi}(1-\tau)^{\varphi} w^{\varphi} y_h^{\varphi} 1_{e_h=e}, \tag{32}$$

$$c_h + a_h \leq (1+r) \sum_{\tilde{h}\in\mathcal{H}} \Pi_{h\tilde{h}}^a \frac{S_{\tilde{h}}}{S_h} a_{\tilde{h}} + \left((1-\tau)1_{e_h=e} + \phi 1_{e_h=u}\right) l_h y_h w, \tag{33}$$

$$S_h = \sum_{\tilde{h}\in\mathcal{H}} \Pi_{h\tilde{h}}^S S_{\tilde{h}}. \tag{34}$$

The steady-state equilibrium is further characterized by the following aggregate equations: market clearing (25)–(27), UI scheme budget balance (28), and factor prices (4). The important result, stated in the following proposition, is that all individual variables can be identified at the steady-state equilibrium.

Proposition 1 *The variables $(\xi_h^u, \Pi_{h\tilde{h}}^u, \Pi_{\tilde{h}h}^a, \Pi_{\tilde{h}h}^S)_{\tilde{h},h\in\mathcal{H}}$ can be identified and computed at the steady-state.*

The proof of this proposition is straightforward. As in the steady-state, we can characterize the stationary wealth distribution of the model, as for any Bewley model. From this stationary distribution, we can integrate policy rules and transition probabilities to determine all individual variables. The individual variables can then be combined to compute the quantities in terms of history bins. In particular, it is possible to compute the transition elements $(\Pi_{h\tilde{h}}^u, \Pi_{\tilde{h}h}^a, \Pi_{\tilde{h}h}^S)_{\tilde{h},h\in\mathcal{H}}$ – which matter for computing bin sizes, projected budget constraints, and projected Euler constraints – as well as the corrective factor $(\xi_h^u)_{h\in\mathcal{H}}$. In addition, we can also identify the set of credit-constrained bins, $\mathcal{C} \subset \mathcal{H}$.

We conclude this section by a convergence result of allocations when the partition $\mathcal{H}$ becomes increasingly fine.

Proposition 2 (Convergence of allocations) *Let $(\mathcal{H}_n)_{n\geq 0}$ be a sequence of partitions, such that: (i) $\mathcal{H}_{n+1}$ is a refinement of $\mathcal{H}_n$ for all n and (ii) the size of partition elements converges to zero. For any n, we denote by $(c_{h_n,t}, a_{h_n,t}, l_{h_n,t})_{h_n}$ the allocations – consumption, savings, and labor supply – associated with partition $\mathcal{H}_n$. We have the following convergence result for allocations:*

$$(c_{h_n,t}, a_{h_n,t}, l_{h_n,t})_{h_n} \longrightarrow_n (c_t(s^\infty), a_t(s^\infty), l_t(s^\infty))_{s^\infty\in\mathcal{S}^\infty},$$

where $(c_t(s^\infty), a_t(s^\infty), l_t(s^\infty))_{s^\infty\in\mathcal{S}^\infty}$ are the allocations of the Bewley model.

This proposition states that the bin allocations converge to the Bewley allocation when the partition becomes increasingly fine. Proposition 2 is general and valid for any type of partition. The proof can be found in Section B of the Appendix. The intuition for the proof is as follows. The projected variable $X_{h,t}$, defined in equation (20), is an average of the value X over a bin h. When the size of this bin converges to zero, the average converges to the "derivative" of the integral, which is in this case the value of X for a unique history. This is, in a sense, a generalized version of the fundamental theorem of calculus.

It is noteworthy that aggregate quantities (capital, total labor supply, output) and prices are by construction equal to their counterparts in the Bewley economy for *any* partition. This still holds at the limit. So, a consequence of Proposition 2 is that, at the steady–state, the projected economy converges to the Bewley economy.

As a result of Proposition 2, we can state the following corollary, stating that all the correcting coefficients vanish when the partition $\mathcal{H}$ becomes increasingly fine.

Corollary 1 (Convergence of correcting coefficients) *Let $(\mathcal{H}_n)_{n\geq 0}$ be a sequence of partitions as in Proposition 2. For any n, we denote by $(\Pi^a_{h_n\tilde{h}_n,t}, \Pi^u_{h_n\tilde{h}_n,t}, \xi^u_{h_n,t})_{h_n,\tilde{h}_n}$ the correcting coefficients – probabilities and concavity corrections – associated with partition $\mathcal{H}_n$. Similarly, $(\Pi^S_{h_n\tilde{h}_n,t})_{h_n,\tilde{h}_n}$ are the transition probabilities for agents' flows. We have the following convergence results:*

$$\xi^u_{h_n,t} \longrightarrow_n 1,\ \Pi^a_{h_n\tilde{h}_n,t} - \Pi^S_{h_n\tilde{h}_n,t} \longrightarrow_n 0,\ \Pi^u_{h_n\tilde{h}_n,t} - \Pi^S_{h_n\tilde{h}_n,t} \longrightarrow_n 0.$$

In words, the difference between the probabilities for agents flows and other correcting probabilities converges to zero. Furthermore, the concavity correction coefficients $(\xi^u_{h_n,t})$ vanish. This means that the factors of the projected model introduce a correction that diminishes when the partition becomes increasingly fine. We will illustrate this convergence numerically in our quantitative exercise of Section 6.

4.2 The dynamics of the approximated model

We now formally present the structure of the approximated model. Following the notation of equation (18), the exact transition probabilities can be written as (for $X = a, u, S$):

$$\Pi^X_{\tilde{h}h,t} = \widetilde{\Pi}^X_{\tilde{h}h,t}\Pi_{y_{\tilde{h}}y_h}\Pi_{e_{\tilde{h}}e_h,t}, \tag{35}$$

where the transition probabilities for productivity, $\Pi_{y_{\tilde{h}} y_h}$ are not time-varying in our economy. Our main assumptions regarding the projected model are stated below.

Assumption A *1. The quantities* $\left(\xi^u_h, \widetilde{\Pi}^u_{\tilde{h},h}, \widetilde{\Pi}^a_{\tilde{h},h}, \widetilde{\Pi}^S_{\tilde{h},h}\right)_{\tilde{h},h\in\mathcal{H}}$ *remain constant and equal to their steady-state values.*

2. for any $h \in \mathcal{H}$, $a_{h,t} = -\bar{a}$ for all t if and only if $a_h = -\bar{a}$.

Assumption A enables us to use our bin-history representation to determine the model with aggregate shocks. As already explained, the first item means that the model still features bin heterogeneity, but that this heterogeneity is not time-varying. However, transition probabilities $(\Pi^u_{\tilde{h}h,t}, \Pi^a_{\tilde{h}h,t}, \Pi^S_{\tilde{h}h,t})_{\tilde{h},h\in\mathcal{H}}$, defined as in equation (35), can be time-varying in the presence of aggregate shocks because job market transition probabilities are also time-varying. In the remainder of the paper, the transitions $(\Pi^u_{\tilde{h}h,t}, \Pi^a_{\tilde{h}h,t}, \Pi_{\tilde{h}h,t})_{\tilde{h},h\in\mathcal{H}}$ should be understood as being time-varying, and constructed using Assumption A.

The second item assumes that if a bin $h \in \mathcal{H}$ is credit-constrained at the steady-state, then it remains credit constrained in the dynamic version of the model. Symmetrically, unconstrained bins at the steady-state are also unconstrained in the dynamic version. This assumption is consistent with the use of a perturbation method, which relies on small aggregate shocks. However, in the general case, the number of credit-constrained households can be time-varying, since the size of the bin can be time-varying.[9] $\mathcal{C} \subset \mathcal{H}$ denotes the set of credit-constrained bins.

We can now formally present our model in the presence of aggregate shocks.

Definition 2 (Model with aggregate shocks) *The approximated model is defined by the following set of equations:*[10]

$$\forall h \in \mathcal{H}\setminus\mathcal{C},\ \xi^u_h U_c(c_{h,t}, l_{h,t}) - \beta\mathbb{E}(1+r_{t+1}) \sum_{\tilde{h}\in\mathcal{H}} \Pi^u_{h\tilde{h},t+1} \xi^u_{\tilde{h}} U_c(c_{\tilde{h},t+1}, l_{\tilde{h},t+1}) = 0, \tag{36}$$

$$\forall h \in \mathcal{C},\ a_{h,t} = -\bar{a}, \tag{37}$$

$$\forall h \in \mathcal{H},\ l_{h,t} = \chi^\varphi (1-\tau_t)^\varphi w_t^\varphi y_h^\varphi 1_{e_h=e}, \tag{38}$$

$$\forall h \in \mathcal{H},\ c_{h,t} + a_{h,t} \leq (1+r_t) \sum_{\tilde{h}\in\mathcal{H}} \Pi^a_{\tilde{h}h,t} \frac{S_{\tilde{h},t}}{S_{h,t}} a_{\tilde{h},t-1} \tag{39}$$

$$+ \left((1-\tau_t)1_{e_h=e} + \phi_t 1_{e_h=u}\right) l_{h,e,t} y_h w_t,$$

[9] It may be possible to relax the second point of Assumption A for modeling large aggregate shocks by using bin-specific penalty functions. We leave this development for future work.

[10] In equation (39) below, $l_{h,e,t}$ is the labor supply of an employed agent with productivity y_h, which determines the UI benefits of unemployed agents of bin h.

together with equations (4) for factor prices, (25)–(27) for market clearing conditions, and (28) for UI scheme budget balance.

Using Definition 2, we can easily simulate the model using standard perturbation methods. Although our goal is to solve for optimal policies with aggregate shocks, an additional advantage of the projection theory is that it improves current simulation techniques. We discuss quantitative comparisons in Section 6. We begin by clarifying how our history-based projection method compares with the method of Reiter (2009).

4.3 Following wealth or histories? Comparison with Reiter's representation

Reiter (2009) develops a method that projects the wealth distribution on a finite set of bins. As with our method, Reiter's algorithm begins by defining bins of wealth from the steady-state distribution. Since the boundaries of the wealth bins are kept constant in the dynamic version and the within-bin wealth distribution is assumed to be uniform, the distribution of wealth is assumed to be a histogram. The transition of agents across bins is computed using agents' saving decisions averaged over each bin. We identify three main differences between Reiter's algorithm and the one examined here.

The first difference is that we construct an explicit approximated model, using budget constraints and Euler equations. As a consequence, we are able to define a consistent Ramsey program for the approximated model with aggregate shocks, as will be performed in Section 5 below. The second difference concerns the simulations: we use more steady-state information and, in particular, we account for the within-bin distribution to construct our model. This improves accuracy when a small number of bins are used.[11]

The third difference relates to the model representation. Reiter's approach is formulated in terms of wealth bins, while our method tracks history bins. To better understand the difference between the "following wealth" and "following histories" approaches, we consider a special case, where TFP is affected by aggregate shocks, but where labor market transitions are constant. In Reiter (2009)'s representation, the measure of agents within each wealth bin changes over time (as the wealth distribution evolves), even though the per capita wealth within each bin remains the same (by construction). In our representation, because of constant transition probabilities across employment statuses, the share of agents within each history bin is constant. Indeed, from

[11] Our construction can be thought of as an efficient way of reproducing the wealth distribution on a finite-dimensional state space. See Algan, Allais, and Den Hann (2010) and Winberry (2018) for a different strategy for approximating the distribution of wealth.

equations (19) and (35), we have $S_h = \sum_{\tilde{h}\in\mathcal{H}} \widetilde{\Pi}^S_{\tilde{h}h} \Pi^S_{e_{\tilde{h}}e_h} S_{\tilde{h}}$, which is constant. The per capita wealth in a given history bin is time-varying, however, because TFP affects consumption and saving choices. To summarize, in Reiter (2009), for any bin, the per capita wealth is constant, while the size is time-varying. Conversely, in our approach, the size of history bins is constant, while the per capita wealth is time-varying.

5 Ramsey program

The previous construction provides a solid foundation for solving Ramsey policies with aggregate shocks. First, computing Ramsey policies in the general case is a very difficult task. It is necessary to introduce additional state variables, such as Lagrange multipliers, for the relevant individual constraints. The Ramsey problem thus involves a joint distribution of two individual state variables – namely, in our case, beginning-of-period wealth and Lagrange multipliers. Characterizing this joint distribution is particularly difficult and, to the best of our knowledge, there is no general method for such a characterization.[12] In our approach, the state-space has finite support, which allows for the resolution of the Ramsey program – using the tools of Marcet and Marimon (2011), for instance. An additional benefit is that it is possible to derive analytical expressions for first-order conditions of the Ramsey program. This eases the interpretation of results and the comparison with other approaches.

5.1 Formulation of the Ramsey program

The Ramsey problem involves determining the unemployment insurance policy (which consists here of the unemployment benefit rate ϕ_t and the labor tax rate τ_t) that corresponds to the "best" competitive equilibrium, according to a utilitarian welfare criterion. To be consistent with the projection, aggregate welfare is measured as $\sum_{t=0}^{\infty} \beta^t \sum_{h\in\mathcal{H}} S_{h,t}\xi^U_{h,t} U(c_{h,t}, l_{h,t})$, where $\xi^U_{h,t}$ are correcting coefficients similar to $\xi^u_{h,t}$ defined in equation (21). These coefficients are related to the non-linearity of the utility function U and are designed to capture the steady-state heterogeneity in history bins. Regarding the coefficients $\xi^U_{h,t}$, we make an assumption similar to Assumption A and it is assumed that the coefficients $(\xi^U_{h,t})_{h\in\mathcal{H}}$ remain constant and equal to their steady-state values, denoted by $(\xi^U_h)_{h\in\mathcal{H}}$ and computed using the steady-state

[12] Açikgöz (2015) proposes an algorithm in an economy without aggregate shocks to approximate the joint distribution. This algorithm is used in Açikgöz, Hagedorn, Holter, and Wang (2018). It relies on certain assumptions about functional forms to find a fixed point.

wealth distribution of the Bewley model. The Ramsey problem can be written as follows:

$$\max_{\left((a_{h,t},c_{h,t},l_{h,t})_{h\in\mathcal{H}},\phi_t,\tau_t\right)_{t\geq 0}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \sum_{h\in\mathcal{H}} S_{h,t}\xi_h^U U(c_{h,t}, l_{h,t})\right], \tag{40}$$

subject to: (i) the budget constraints (24), (ii) the labor Euler equations (29), (iii) the consumption Euler equations (30), (iv) the UI scheme budget balance (28), (v) the market clearing constraints (25) and (27), and finally (vi) the factor prices w_t and r_t (4). Note that we also have to take into account the constraints driving the evolution of the bin sizes (19). However, since the evolution is independent of the planner's choices, it has no impact on the Ramsey policies.

A reformulation of the Ramsey problem. We simplify the formulation of the Ramsey problem exposed in equation (40). We first denote by $\beta^t S_{h,t}\lambda_{h,t}$ the Lagrange multiplier of the Euler equation for history bin h at date t, (36). These Lagrange multipliers are key to understanding the planner's program. If agents' private incentives to save in bins h at date t are socially optimal, then their Euler equation is not a constraint and the Lagrange multiplier is 0: $\lambda_{h,t} = 0$. Then, depending on how the planner perceives the distortions of saving incentives (i.e., whether agents save too much or too little from the planner's perspective), these coefficients can be either positive or negative.[13]

We also define for all $h \in \mathcal{H}$:

$$\Lambda_{h,t} \equiv \frac{\sum_{\tilde{h}\in\mathcal{H}} S_{\tilde{h},t}\Pi^S_{\tilde{h}h,t}\lambda_{\tilde{h},t-1}}{S_{h,t}}, \tag{41}$$

which, for agents in history bin h, can be interpreted as the average of their previous period Lagrange multipliers for the Euler equation. Finally, note that $\lambda_{h,t} = 0$ if $a_{h,t} = -\overline{a}$: the multiplier $\lambda_{h,t}$ is null when the credit constraint is binding. The product $\lambda_{h,t}\nu_{h,t}$ (for any t and h) is thus always null. This property is key for simplifying the expressions. The following lemma summarizes our simplified Ramsey problem.

[13]We provide a simple example in the Technical Appendix to clarify the relationship between saving incentives and the sign of the Lagrange multiplier for Euler equations.

Lemma 1 (Simplified Ramsey problem) *The Ramsey problem can be simplified into:*

$$\max_{\left((a_{h,t},c_{h,t},l_{h,t})_{h\in\mathcal{H}},\phi_t,\tau_t\right)_{t\geq 0}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{h\in\mathcal{H}} S_{h,t} \Big(\xi_h^U U(c_{h,t},l_{h,t}) \tag{42}$$

$$- \left(\lambda_{h,t} - (1+r_t)\Lambda_{h,t}\right) \xi_h^u U_c(c_{h,t},l_{h,t})\Big)$$

$$s.t.\ \ \lambda_{h,t} = 0 \ \mathit{if}\ a_{h,t} = -\overline{a}, \tag{43}$$

and subject to equations (24), (25), (27), (28), (29), and (4).

The contribution of the Lemma is to show that the factorization of the Lagrangian can be easily performed with partitions in the space of histories, as shown by Marcet and Marimon (2011) for agents. This considerably simplifies the derivation of first-order conditions of the Ramsey program. The proof, provided in the Technical Appendix, is based on a re-writing of the Lagrangian to introduce Lagrange multipliers into the objective.

5.2 Ramsey conditions and economic interpretation

Using proper substitution, the program (42)–(43) can be written as a maximization problem with only two sets of choice variables: the labor tax τ_t and saving choices $(a_{h,t})_{h\in\mathcal{H},t\geq 0}$. The current section derives the planner's first-order conditions for any general partition and discusses the economic trade-offs determining the time-varying replacement rate.

To ease the economic interpretation of the first-order conditions, we define the following useful aggregates:

$$\Psi_{h,t} = \xi_h^U U_{c,h,t} - \left(\lambda_{h,t} - (1+r_t)\Lambda_{h,t}\right) \xi_h^u U_{cc,h,t}, \tag{44}$$

where $U_{c,h,t} = U_c(c_{h,t},l_{h,t})$ is the consumption marginal utility and $U_{cc,h,t} = U_{cc}(c_{h,t},l_{h,t})$ is the derivative of the marginal utility. The quantity $\Psi_{h,t}$ will be called the *marginal social valuation of liquidity* for agents in history bin h, because it is the marginal gain for the planner of transferring resources in bin h at date t. If an agent receives one additional unit of goods today, this additional unit will have a value proportional to $U_{c,h,t}$. This value only accounts for private valuation, but should also reflect, for the planner, the effect on the saving incentives, i.e., on the Euler equations. This additional unit therefore affects the agent's saving incentive from period $t-1$ to period t and from period t to period $t+1$. This effect is captured by the second term, which is proportional to $U_{cc,h,t}$.

The saving decision in bin h at date t affects the individual welfare of all agents due to general equilibrium effects on capital and prices. The first-order condition of the Ramsey program related to saving choices summarizes all these effects. It can be written as follows for unconstrained bins $h \in \mathcal{H} \setminus \mathcal{C}$:

$$\Psi_{h,t} = \underbrace{\beta \sum_{\tilde{h}\in\mathcal{H}} \mathbb{E}_t \left[(1+r_{t+1})\Pi^a_{h\tilde{h},t}\Psi_{\tilde{h},t+1}\right]}_{\text{liquidity smoothing}} + \beta \frac{\alpha(1-\alpha)}{1+\alpha\varphi} \frac{1}{L_{t+1}} \left(\frac{K_t}{L_{t+1}}\right)^{\alpha-1} \tag{45}$$
$$\times \left(\underbrace{\mathbb{E}_t \left[(1-\tau_{t+1}) \sum_{\tilde{h}} \Psi_{\tilde{h},t+1} S_{\tilde{h},t+1} l_{\tilde{h},t+1} \tilde{y} 1_{e_{\tilde{h}}=e}\right]}_{\text{net wage effect for employed}} \right.$$
$$+ (1+\varphi) \underbrace{\mathbb{E}_t \left[\phi_{t+1} \sum_{\tilde{h}} \Psi_{\tilde{h},t+1} S_{\tilde{h},t+1} l_{\tilde{h},e,t+1} \tilde{y} 1_{e_{\tilde{h}}=u}\right]}_{\text{wage effect on unemployment benefits for unemployed}}$$
$$\left. - \left(\frac{K_t}{L_{t+1}}\right)^{-1} \sum_{\tilde{h}\in\mathcal{H}} \mathbb{E}_t \underbrace{\left[S_{\tilde{h},t+1}\left(\Lambda_{\tilde{h},t+1}\xi^u_{\tilde{h}} U_{c,\tilde{h},t+1} + \Psi_{\tilde{h},t+1}\tilde{a}_{\tilde{h},t+1}\right)\right]}_{\text{interest rate effect on smoothing and wealth}} \right),$$

where $\tilde{a}_{h,t} = \sum_{\tilde{h}\in\mathcal{H}} \Pi^a_{\tilde{h}h,t} \frac{S_{\tilde{h},t-1}}{S_{h,t}} a_{\tilde{h},t-1}$ is the beginning-of-period wealth in bin h at date t[14].

Equation (45) features the first-order condition on the liquidity allocation (i.e., saving choices) for unconstrained bins. The equation, though apparently complicated, has a straightforward interpretation. Four effects are at play. The first one is a direct effect that measures the expected future value of liquidity tomorrow. In other words, this component states that liquidity value should be smoothed over time. This first part is very similar to a standard Euler equation. We refer to this first term as "liquidity smoothing". The three other components alter the pure smoothing effect and reflect the fact that the planner also takes into account the consequences of liquidity allocation on prices. More precisely, the second and third components correspond to the marginal effect of additional saving on the wage rate. This affects employed agents (second component) and unemployed (third component) agents, because UI benefits are proportional to the labor income of employed agents with the same productivity. Finally, the fourth and last component reflects the distortions of the interest rate on smoothing incentives and on wealth accumulation.

[14] Compared to Marcet and Marimon (2011) or Chien, Cole, and Lustig (2011), we don't use cumulative Lagrange multipliers to analyze the dynamics. Instead, we use the period multipliers to derive first-order conditions of the planner. Indeed, these conditions are easier the interpret with period multipliers and the simulation of the model relies on a smaller number of variables. See these two references for a discussion of the existence of these multipliers in such economies.

The second first-order condition, relating to the labor tax, can be written as follows:

$$\underbrace{\frac{\alpha}{(1-\tau_t)K_{t-1}}\sum_{\tilde{h}\in\mathcal{H}} S_{\tilde{h},t}\left(\Lambda_{\tilde{h},t}\xi^u_{\tilde{h}}U_{c,\tilde{h},t}+\Psi_{\tilde{h},t}\tilde{a}_{\tilde{h},t}\right)}_{\text{effect on prices, smoothing and redistribution}}+\frac{1}{\varphi}\underbrace{\sum_{\tilde{h}\in\mathcal{H}} S_{\tilde{h},t}\Psi_{\tilde{h},t}\frac{l_{\tilde{h},t}}{L_t}\tilde{y}1_{e_{\tilde{h}}=e}}_{\text{cost of the tax for employed}} = \underbrace{\frac{S_{e,t}}{S_{u,t}}\left(\frac{1}{\varphi}+1-\frac{1-\alpha}{1-\tau_t}\right)\sum_{\tilde{h}\in\mathcal{H}} S_{\tilde{h},t}\Psi_{\tilde{h},t}\frac{l_{\tilde{h},e,t}}{L_t}\tilde{y}1_{e_{\tilde{h}}=u}}_{\text{gain of unemployment benefits for unemployed}}. \tag{46}$$

Equation (46) determines the optimal labor tax rate by setting the marginal costs of a higher tax rate equal to the marginal benefits. On the left-hand side of (46), marginal costs comprise two effects. The first reflects the tax distortion on the interest rate and thus on saving incentives. The second marginal cost on the left-hand side of (46) reflects the impact of the labor tax on employed agents, taking into account the negative net effect on the labor supply (inversely proportional to the Frisch elasticity φ). On the right-hand side of equation (46), the marginal benefit comprises the marginal gain of tax (and UI benefit) for unemployed agents. Finally, equation (46) embeds, in a compact form, the general equilibrium effect on wages, which are captured by both the Frisch elasticity of the labor supply φ, and the concavity of the production function α.

5.3 Calibration and simulation of the Ramsey allocation using the approximated model

It is straightforward to derive the Ramsey conditions to be applied to the approximated model, as there are a finite number of equations. In this environment, standard perturbation methods can be used to simulate the model with aggregate shocks. However, an important first step is to obtain the right steady-state allocation, and thus a consistent approximated model. The general method is the following:

1. Solve the "true" Bewley model (i.e., without aggregate shocks) for a given UI policy.

2. Construct the projected and approximated model, and derive implied values of the Lagrange multipliers, using (45).

3. Iterate on the UI policy until the optimality condition (46) is satisfied.

This strategy has three advantages. First, the derivation of correcting parameters for the approximated model is consistent with the true Bewley model at each step. Second, since the "true" Bewley model is required to exist at each step of the iteration, the perturbation method is not used around non-existing, steady-state equilibria.[15] A final advantage is that the steady-state value of Lagrange multipliers can be expressed in closed-form using matrix calculus, as the state-space has a finite dimension. This provides a very efficient algorithm with which to compute the steady-state solution of the Ramsey program, as shown in Appendix E.

6 Numerical analysis

We perform two separate numerical exercises. First, in Section 6.1, we solve a simple model with an exogenous replacement rate to compare our projection method with existing solution techniques. Second, in Section 6.2, we use the projection method to solve for the optimal replacement rate in the general model presented above.

6.1 Comparing solution methods: Krusell-Smith, explicit and implicit partitions

We perform a similar exercise to den Haan (2010) in order to compare our solution methods to the Krusell-Smith method (henceforth KS). The economy under consideration features only unemployment risk, with a constant and exogenous replacement rate, and each employed worker inelastically supplies one unit of labor. This economy is thus a special case of the economy presented in the paper.

The period is a quarter. The calibration is standard and based on den Haan (2010). For the discount factor, the capital share, the depreciation rate, the credit limit, and the replacement rate, we have $(\beta, \alpha, \delta, \bar{a}, \phi) = (0.99, 0.36, 0.025, 0, 0.1)$. The economy can be in one of two aggregate states, G (standing for Good) or B (Bad). In the Good aggregate state, the productivity level is $z_G = +1\%$ and the unemployment rate is $U_G = 4\%$. In the Bad aggregate state, the productivity level is $z_B = -1\%$ and the unemployment rate is 10%. The transition probabilities on the labor market depend on the aggregate state. The transition matrix is given in Table 1. The probability of transitioning from unemployment in the Bad state $(B, 0)$ to employment in the Good state $(G, 1)$ is 9.3750%, for instance.

We solve the previous model using the KS algorithm refined by Maliar, Maliar, and Valli (2010). Their refinement is based on an endogenous grid and on the iteration over the aggregate

[15] In the general case, it is possible that the set of equilibria is larger in the approximated model than in the true Bewley one. This procedure ensures that we only select equilibria for the original Bewley model.

	$B, 0$	$B, 1$	$G, 0$	$G, 1$
$B, 0$	52.5000	35.0000	3.1250	9.3750
$B, 1$	3.8889	83.6111	0.2083	12.2917
$G, 0$	9.3750	3.1250	29.1667	58.3330
$G, 1$	0.9115	11.5885	2.4306	85.0694

Table 1: Transition probabilities on the labor market (expressed in %).

law of motion of capital stock. Here, we follow their procedure exactly and use $1,000$ grid points and simulate $10,000$ agents for $10,000$ periods. This type of algorithm is known to provide accurate results in this simple environment (see den Haan 2010).

Constructing projected models. To simulate the model with history-representation using either explicit or implicit partitions, we first need to solve the model steady-state. Steady-state job market transition probabilities can be deduced from Table 1.

As we use the perturbation method when simulating the model with history-representation, we approximate the discrete aggregate risk process described above with its continuous counterpart. The log of TFP, $z_t = \log(Z_t)$, follows an AR(1) process $z_t = 0.95 z_{t-1} + \varepsilon_t$ where (ε_t) are iid Gaussian white noises with zero mean and a standard deviation $\sigma_z = 0.66\%$. We use the following process for the unemployment rate and transition probabilities:

$$\Pi_{ue,t} = 0.5257 + 10.2978 z_t + 5.9944 z_{t-1}, \tag{47}$$

$$S_{u,t} = 7\% - 3 z_t. \tag{48}$$

We obtain this calibration by replicating the first- and second-order moments (including correlations) of the discrete processes in Krusell and Smith (1998).

To construct the projected model, we first solve the Bewley model using value function iteration to directly obtain policy rules on the relevant grid, which simplifies the projection. We use an exponential grid with $1,000$ points. We then project the model for both explicit and implicit partitions. In the first case, we truncate the idiosyncratic histories with $N^{ex} = 6$, such that there are 64 bins. For implicit partitions, we use $N^{im} = 10$, such that there are 20 bins. These partitions provide satisfactory dynamics results as shown below. The simulations with aggregate shocks are based on a linear approximation of the model – computed with Dynare – which is known to generate accurate results (as discussed recently by Boppart, Krusell, and

Mitman 2018).

Comparing the steady-state distribution of wealth for implicit and explicit partitions We first report steady-state distributions for the Bewley model and for the projected model on implicit and explicit partitions, as shown in Table 2.

Decile	1	2	3	4	5	6	7	8	9	10
Bewley model	3.54	9.12	16.61	25.82	36.44	48.01	60.44	73.46	86.73	100
Implicit part.	3.56	9.14	16.63	25.87	36.45	48.05	60.44	73.46	86.73	100
Explicit part.	8.48	18.00	28.10	38.37	48.64	58.91	69.18	79.46	89.73	100

Table 2: Cumulative distribution of wealth by decile for the Bewley model and for the implicit-partition ($N^{im} = 10$) and explicit-partition models ($N^{ex} = 6$).

The implicit-partition model reproduces almost exactly the Bewley model wealth distribution. This is almost a feature of the construction, as this information is used to calibrate the model. The explicit partition is less efficient at reproducing the Bewley model wealth distribution. The reason is the large bin of employed agents for $N^{ex} = 6$ consecutive periods, which contains 70% of the population. This is due to the high persistence of the employment state.

6.1.1 Results

We now compare the outcomes of three different algorithms: the KS algorithm, the projection with implicit partitions, and the projection with explicit partitions. For each of the three algorithms, we compare the moments generated by simulations of the model for $10,000$ periods with those implied by theory, in the case of both implicit and explicit partitions. The ability to compute theoretical moments is an additional positive aspect of our solution technique, which allows us to quantify sampling errors for endogenous variables.

Table 3 reports first- and second-order moments for three aggregate variables: output (Y), consumption (C), and capital (K), and for three labor market quantities: unemployed population share S_u, job finding rate Π_{ue}, and job separation rate Π_{eu}. It also contains the computational time needed to solve the model with aggregate shocks.[16] Economy (1) presents the results for the KS method, for the simulated economy. Economy (2) presents the theoretical moments

[16]For economies 3-6, we report the time to perform the projection and the perturbation once the Bewley model is solved.

for the same economy for exogenous variables only, as theoretical moments for endogenous variables cannot be computed. Economies (3) and (4) present the results for the explicit partition, for simulated and theoretical moments, respectively. Economies (5) and (6) present the same moments for implicit partition.

Methods		Krusell-Smith		Explicit ($N^{ex} = 6$)		Implicit ($N^{im} = 10$)	
Economies		(1)	(2)	(3)	(4)	(5)	(6)
Moments (in)		Simul.	Theory	Simul.	Theory	Simul.	Theory
Y	mean	3.4469	-	3.4525	3.4525	3.4525	3.4525
	std	0.1184	-	0.1179	0.1175	0.1187	0.1178
C	mean	2.5605	-	2.5637	2.5637	2.5637	2.5637
	std	0.0459	-	0.0461	0.0482	0.0464	0.0478
K	mean	35.4626	-	35.5509	35.5509	35.5509	35.5509
	std	0.8788	-	0.8027	0.7967	0.8501	0.8748
L	mean	0.9292	0.9300	0.9300	0.9300	0.9300	0.9300
	std	0.0300	0.0300	0.0302	0.0300	0.0302	0.0300
Π_{eu}	mean	0.0374	0.0372	0.0376	0.0376	0.0376	0.0376
	std	0.0129	0.0130	0.0160	0.0162	0.0160	0.0162
Π_{ue}	mean	0.5257	0.5292	0.5000	0.5000	0.5000	0.5000
	std	0.1529	0.1533	0.1548	0.1533	0.1548	0.1533
Speed (in s.)		705	-	3.3	2.5	1.7	1.4

Table 3: Comparing moments with different resolution techniques.

The results are very similar for both first-order and second-order moments. For second-order moments, the sampling error is small but not negligible. For instance, the sampling error for the standard deviation of output (line Y, std) is 9.10^{-4} for the implicit-partition method (absolute difference between Economies 5 and 6). This difference is greater than the difference between the KS and the implicit-partition method (absolute difference between Economies 1 and 5 for the line Y, std equals 3.10^{-4}). The implicit-partition solution is slightly more accurate (and faster) than the explicit-partition solution. Furthermore, second-order moments are close to those of the KS economy, notably for the capital stock. From this comparison, we conclude that our projection theory performs well compared to the global method and is faster.

6.2 Optimal unemployment benefits at the steady-state

We now solve for the optimal unemployment benefits in the general model with both employment and productivity risks. The period is a quarter. The calibration is adapted from Krueger, Mittman, and Perri (2018). As before, the capital share is $\alpha = 0.36$ and the depreciation rate is $\delta = 0.025$. The TFP process is a standard AR(1) process for TFP shocks, $Z_t = \exp(z_t)$, with:[17]

$$z_t = \rho_z z_{t-1} + \varepsilon_t^z, \tag{49}$$

where $\varepsilon_t^z \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_z^2)$. We use the standard values of $\rho_z = 0.95$ and $\sigma_z = 0.31\%$ to obtain a standard deviation of the TFP shock z_t equal to 1% at a quarterly frequency.

The productivity risk is a first-order process estimated from PSID data by Krueger, Mittman, and Perri (2018):

$$\log y_t = \rho_y \log y_{t-1} + \varepsilon_t^y,$$

with $\varepsilon_t^y \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_y^2)$, $\rho_y = 0.9923$, and $\sigma_y = 9.90\%$. The productivity process is discretized, using the Rouwenhorst (1995) procedure, into 7 idiosyncratic states with a constant transition matrix. As agents can be either employed or unemployed, each agent can be in $14 = 7 \times 2$ idiosyncratic states.

For the labor market, we follow Shimer (2003) and assume that the job-separation rate is constant over the business cycle, and that the job-finding rate is time-varying and procyclical. We find $\Pi_{eu}^{SS} = 4.87\%$ for the average job-separation rate, $\Pi_{ue}^{SS} = 78.6\%$ for the average job-finding rate, and 10% for the standard deviation of the job-finding rate.[18] As the standard deviation of z_t is 1%, we assume that the job-finding rate is defined as follows:

$$\Pi_{ue,t} = \Pi_{ue}^{SS} + \sigma_{ue} z_t, \text{ with } \sigma_{ue} = 10.$$

The period utility function is $u(C) = \ln(C)$ and the discount factor is equal to $\beta = 0.99$. Various values of the Frisch elasticity of labor supply, φ, are used in quantitative work (see Chetty, Guren, Manoli, and Weber 2011). Following Heathcote (2005) we use $\varphi = 0.3$, which is in the lower range of empirical estimates, but may be more realistic for business cycle dynamics. The scaling parameter is set to $\chi = 0.38$, which implies normalizing the aggregate labor supply

[17] Krueger, Mittman, and Perri (2018) use a two-state Markov process representation of this AR(1) process in order to use the KS algorithm.

[18] See the estimates of Challe and Ragot (2016), for instance. Krueger, Mittman, and Perri (2018) estimate a labor process to match labor market dynamics in normal times and in severe recessions. Their labor process estimation appears to be close to ours, with an implied autocorrelation of $\Pi_{ue,t}$ of 0.96 and a standard deviation of 0.4%.

to 1. Table 4 provides a summary of the model parameters.

Parameter	Description	Value
β	Discount factor	0.99
α	Capital share	0.36
δ	Depreciation rate	0.025
Π_{ue}^{SS}	Average job finding rate	78.6%
$\bar{a}$	Credit limit	0
Π_{eu}^{SS}	Average job separation rate	4.87%
U	Steady-state unemployment rate	5.83%
ρ_z	Autocorrelation TFP	0.95
σ_z	Standard deviation TFP shock	0.31%
σ_{ue}	Cov. job find. rate with TFP	10
ρ_y	Autocorrelation idio. income	0.9923
σ_y	Standard dev. idio. income	9.90%
χ	Scaling param. labor supply	0.38
φ	Frisch elasticity labor supply	0.3

Table 4: Parameter values of the baseline calibration. See the text for descriptions and targets.

To ease the analysis, the list of the model equations can be found in Appendix D. Because of the 14 idiosyncratic shocks, we use an implicit partition. An explicit partition would generate a state-space too large for this high number of idiosyncratic states. For each idiosyncratic state, we consider 15 equally-populated history bins. Accordingly, we follow $14 \times 15 = 210$ agent bins in the dynamic model. We solve the dynamic model using the Dynare solver, to generate a first-order perturbation around the steady-state. The Ramsey allocation is determined by $1,520$ equations, including the planner's first-order conditions. The algorithm presented in Appendix F solves the model in 40 seconds.

6.2.1 Results

The optimal replacement rate is found to be $\phi^{SS} = 79\%$ at the steady-state for the baseline calibration presented in Table 4. As mentioned above, the current model does not consider other distortions studied in the literature on UI (relating to moral hazard) and does not include other public spending finance needs. This evaluation should thus be considered as an upper estimate.

Before analyzing the quantitative trade-offs generating this replacement rate, we first analyze the convergence properties of the model when we change the structure of the partition. The first four lines of Table 5 report the convergence properties of the model's correcting parameters ξ^U and ξ^u when the number of elements N in the wealth partition increase, but when all history bins remain of equal size. ξ^U corrects the within-bin heterogeneity for the utility function U, while ξ^u corrects the heterogeneity for the marginal utility of consumption appearing in the Euler equations. As expected, the between-bin variance of each of these quantities, denoted by $var(\xi^U)$ and $var(\xi^u)$ respectively, decreases monotonically with the number of elements, N. The choice $N = 15$ appears to be a good trade-off between the inequality representation and the accuracy of the aggregate dynamics. The corresponding Gini for wealth is 0.78 in the projected economy for $N = 15$, which is the same value as in the Bewley economy. The last two lines of Table 5 analyze the convergence properties of the matrices Π^a and Π^u, which correct for heterogeneity in transitions, reporting the standard deviation $std\left(\Pi^S_{i,j} - \Pi^u_{i,j}\right)$ over $i, j = 1...210$. As expected, these variances tend toward 0, and matrices are close and close to Π^S as N increases. Finally, the aggregate dynamics appear not to depend significantly on N, for values of N greater than 10.

N	5	10	15	20
$mean\left(\xi^U\right)$	0.9040	1.0010	0.9990	0.9999
$std(\xi^U)$	0.7166	0.0384	0.0097	0.0087
$mean\left(\xi^u\right)$	1.0259	1.0120	1.0062	1.0044
$std(\xi^u)$	0.0377	0.0259	0.0160	0.0138
$std\left(\Pi^S - \Pi^u\right)$	0.0625	0.0463	0.0344	0.0303
$std\left(\Pi^S - \Pi^a\right)$	0.0240	0.0151	0.0101	0.0080

Table 5: Convergence properties of correcting coefficients ξ^U, ξ^u ,Π^a, and Π^u for increasing N. See text for details.

To provide further intuitions of the distortions in this economy, Table 6 gathers statistics for the Lagrange multipliers λ of Euler equations. As already explained, these Lagrange multipliers can be positive, if agents are not saving enough, or negative if they are saving too much – as seen from the planner's perspective. We consider the case where the replacement rate is set to $\phi^{SS} = 79\%$ and the case where $\phi = 50\%$ to examine the effects on distortions when the replacement rate is lower. For both cases, we compute the mean and the standard deviation of the Lagrange multipliers between bins, denoted by $mean\,(\lambda)$ and $std\,(\lambda)$, respectively. We also report the correlation $corr(\lambda, \tilde{a})$ between the Lagrange multiplier and beginning-of-period wealth. In the baseline case of $\phi = 79\%$, the average value of Lagrange multipliers between agents is found to be positive. As a consequence, the planner would like agents to save more and the capital stock to increase. A higher capital stock would raise the income of low-productivity agents, as discussed in Dávila, Hong, Krusell, and Ríos-Rull (2012). The standard deviation is found to be large and the correlation between Lagrange multipliers and beginning-of-period wealth is positive. Poor agents have a negative multiplier λ, whereas rich agents have a positive one. The planner would therefore like poor agents to save less, whereas rich agents should save more. The planner cannot reach this outcome by relying solely on the replacement rate, because decreasing the replacement rate increases the capital stock by increasing the saving incentives of wealth-poor agents for self-insurance purposes. These points are confirmed when considering the lower replacement rate of $\phi = 50\%$. The capital stock K is higher than in the case where $\phi = 79\%$. Agents save more and the average value of the Lagrange multiplier λ is smaller. This is obtained by an increase in consumption inequality between employed and unemployed agents (c^u/c^e), which decreases from 88%, when $\phi = 79\%$, to 81% when $\phi = 50\%$.[19]

Replacement rate	$mean\,(\lambda)$	$std(\lambda)$	$corr(\lambda, \tilde{a})$	K	c^u/c^e
$\phi = 79\%$	105	762	0.85	47.79	88%
$\phi = 50\%$	53	306	0.89	48.10	81%

Table 6: Statistics regarding the Lagrange multipliers λ of Euler equations.

Comparison with other welfare criteria It may be useful to compare our steady-state results to those implied by other welfare criteria, which do not solve the planner's program. A first strategy consists in maximizing the steady-state welfare – see for instance Aiyagari and McGrattan (1998) for an early example. This strategy generates very different allocations. With this criterion, we find that the optimal replacement rate is close to 0, which is consistent

[19] It is known that incomplete insurance markets can generate either an over-accumulation or an under-accumulation of capital compared to the constrained efficient equilibrium, see Dávila, Hong, Krusell, and Ríos-Rull (2012) and Aiyagari (1995) for two different cases. The optimal replacement rate is a trade-off between this distortion and the amount of insurance.

with the results of Krusell, Mukoyama, and Sahin (2010). Agents hold high savings for self-insurance motives, which yields a very high capital stock and very high aggregate consumption. Internalizing the cost of accumulating this capital stock – as the planner does in the Ramsey program – generates a higher replacement rate.

Optimal unemployment benefits over the business cycle In Figure 1, we first report the IRFs after a positive aggregate shock for selected variables to provide intuition about the optimal behavior of the replacement rate. The first panel of Figure 1 (labeled dz) plots the TFP IRFs after a positive innovation of TFP of one standard deviation (in percent), which corresponds to a 0.3% increase. The second panel (labeled df) is the job-finding rate. It is procyclical and increases by 3%, which is, by construction, 10 times more than TFP. The job-finding rate thus increases from 78% to 81%. The GDP, plotted in the third panel (dY), increases by 0.5%. Total labor, in efficient units, increases by 0.4%, as reported on the fourth panel (dL). Both aggregate consumption (5th panel, dC) and capital (6th panel, dK) are also increasing, but hump-shaped, which is a standard outcome. The third line of Figure 1 first plots unemployment (dU), which decreases by 0.2%. The second panel ($dphi$) of this third line is the replacement rate, which increases by almost 4% (from 79% to 83%). The replacement rate – and the related labor tax rate – are thus procyclical, which enables the planner to diminish the volatility of the economy (see Table 7 below). The last panel ($ratio\_c$) plots a measure of inequality, which is the average consumption of unemployed agents over the average consumption of employed agents c^u/c^e. This ratio decreases after a positive shock, implying that the ratio of the consumption of unemployed to employed households falls. Inequalities are rising and procyclical in this economy. The higher replacement rate and the higher labor tax do not fully offset the increasing income inequality after a positive TFP shock.

Second-order moments. Table 7 provides second-order moment statistics for the economy with the optimal time-varying replacement rate (first row) and for the economy with a constant replacement rate set at the optimal steady-state value ϕ^{SS} (second row). Obviously, by construction, the first-order moments are identical in both economies. Following the business cycle literature since Cooley and Prescott (1995), we compute the log of variables, except for the replacement rate ϕ, and we HP-filter them (with an HP-filter parameter set to $1,600$).[20]

The first lesson from Table 7 is that a time-varying replacement rate enables the planner to

[20] In the Technical Appendix, we report additional results, such as moments in levels.

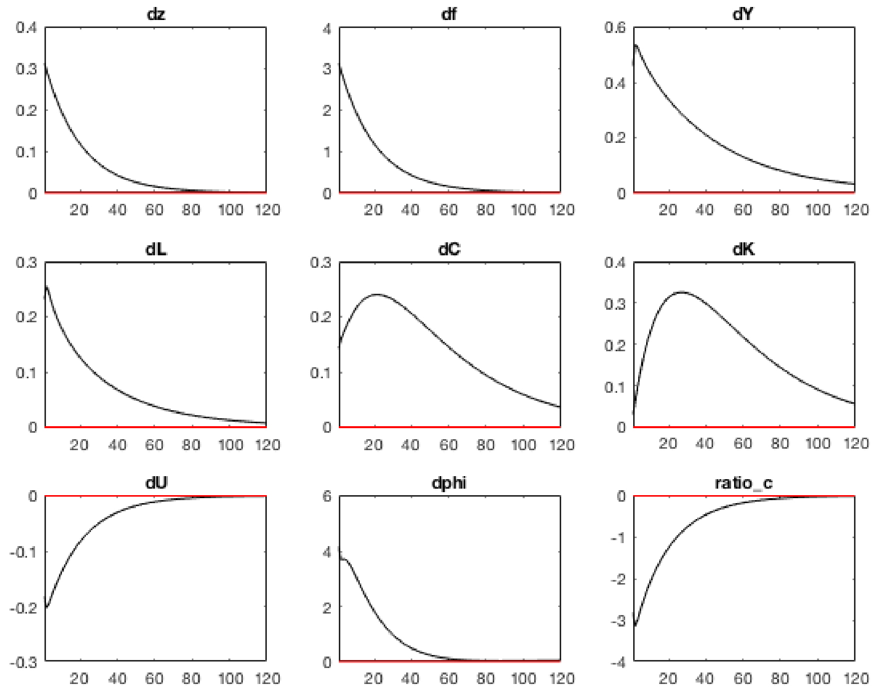

Figure 1: IRFs after a TFP positive shock. The variables $dz, df, dphi$ are in shown in percentage-level deviation from the steady-state. The variables $dY, dC, dK, dL, ratio\_c$ are shown in the percentage proportional deviation from the steady-state value. See the text for a description of the variables.

reduce the volatility of the main aggregate variables – except for the replacement rate, obviously. The reduction is especially important for labor and consumption, and to a lesser extent for output. The standard deviation of consumption falls from 0.28 to 0.20 when the replacement rate is optimally time-varying. The channel comes from a reduction in the volatility of inequality over the business cycle, as can be seen by the reduction in the volatility of ratio_c. The impact on investment is in the same direction as the other effects but the size is one order of magnitude smaller. This reduction in the overall volatility is generated by the pro-cyclical replacement rate – as can be seen from column (Y, ϕ). The replacement rate – and the corresponding labor tax – are high in booms and low in recessions.

Dynamics Property of the optimal replacement rate. As a final investigation, we analyze if a simple rule, relating the optimal replacement rate only to the moments of the distribution of wealth, could provide a good approximation. Doing so do not consider the dynamics of the

	Standard Deviation (%)						Correlations			
	C	I	Y	L	ϕ	ratio_c	(Y, Y_{-1})	(Y, C)	(Y, ϕ)	$(Y, \text{ratio\_c})$
Optimal ϕ	0.20	1.67	0.69	0.30	5.7	4.15	0.83	0.95	0.98	-0.995
$\phi = \phi^{SS}$	0.28	1.69	0.72	0.45	0	4.61	0.80	0.98	0	-0.998

Table 7: Standard deviation (in %) and correlation for key variables and for time-varying ϕ (first line) and constant ϕ (second line). All variables have been logged, except ϕ, and HP-filtered with the parameter set to $1,600$.

distribution of Lagrange multipliers, which are in fact part of the state space. The result of a regression of the replacement rate ϕ on the technology shock z, the first, second and third-order moments of the wealth distribution generates a R^2 of 0.9879, and all variables are significant. This R^2 is actually low in such economies, for which the R^2 is usually closer to 1 to forecast aggregate variables. Indeed, simulating the model for $10,000$ periods, the difference between the optimal replacement rate and the one implied by the result of the regression is as high as 8%. From this experiment, it is unlikely that one can exclude the dynamics of the Lagrange multipliers from the state space to approximate the dynamics of the instruments of the planner.

To conclude, the optimal replacement rate does more than merely stabilize the income ratio between employed and unemployed agents, which would be the case with a constant replacement rate. The optimal replacement rate is procyclical and helps to reduce the difference in the volatility of the consumption ratio of unemployed to employed agents. However, labor tax distortions prevent the optimal policy rate from fully isolating employment risk from aggregate risk. As a consequence, inequalities are procyclical.

7 Conclusion

This paper presents a projection theory of sequential representations of incomplete insurance market models. We use a finite partition of the space of idiosyncratic histories to construct an intuitive approximated model, which can be easily simulated with aggregate shocks, and for which optimal Ramsey policies can be derived. The paper applies the theory to characterize optimal time-varying unemployment benefits when the economy is hit by both technology-related and labor market shocks. The optimal replacement rate is procyclical and helps to reduce consumption volatility.

The simulation of the model uses perturbation methods, which considerably eases implementation. Such methods, however, rely on small aggregate shocks around a well-defined steady-state. They are less relevant for models with large macroeconomic shocks, for which additional developments are needed, using global methods.

There are two main directions for further research in the current framework. The first would be to provide a deeper analysis of the design of optimal partitions in the sequential representation. In the current paper, we use two partitions, a truncation theory and an implicit partition based on wealth distribution. For more complex models, with either a higher number of idiosyncratic states or a wider set of policy tools for the planner, it would be useful to have a general theory of the optimal partition structure, which could be a mix of both partitions, depending on the planner's assigned objective. Second, the theory obviously opens the possibility of considering other applications. The underlying model could be generalized to consider any relevant frictions on the goods, labor, or financial markets, such as limited participation on financial markets or nominal frictions. In addition, the planner could use other tools to reduce distortions, such as a whole set of fiscal instruments or monetary policy instruments. The simplicity of the implementation could contribute to a more systematic integration of redistributive effects in the design of economic policies.

References

Ábrahám, A., J. Brogueira de Sousa, R. Marimon, and L. Mayr (2019): "On the Design of a European Unemployment Insurance System," Working Paper, European University Institute.

Açikgöz, O. (2015): "Transitional Dynamics and Long-Run Optimal Taxation Under Incomplete Markets," Working Paper, Yeshiva University.

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

Adjemian, S., H. Bastani, M. Juillard, F. Karamé, J. Maih, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot (2011): "Dynare: Reference Manual, Version 4," Dynare Working Paper 1, CEPREMAP.

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2017): "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual 2017*, 32.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109(3), 659–684.

——— (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103(6), 1158–1175.

Aiyagari, S. R., and E. R. McGrattan (1998): "The Optimum Quantity of Debt," *Journal of Monetary Economics*, 42(3), 447–469.

Algan, Y., O. Allais, and W. J. Den Hann (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using Parameterized Cross-Sectional Distributions," *Journal of Economic Dynamics and Control*, 34(1), 59–68.

Bewley, T. F. (1983): "A Difficulty with the Optimum Quantity of Money," *Econometrica*, 51(5), 1485–1504.

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.

Bilbiie, F., and X. Ragot (2017): "Monetary Policy and Inequality when Aggregate Demand Depends on Liquidity," CEPR Working paper, Centre for Economic Policy Research.

Billingsley, P. P. (1965): *Ergodic Theory and Information.* John Wiley & Sons, Inc., New Yok.

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.

Brémaud, P. (2014): *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues.* MIT Press, New York, springer-verlag edn.

Challe, E. (2018): "Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy," *American Economic Journal: Macroeconomics*, forthcoming.

Challe, E., F. LeGrand, and X. Ragot (2013): "Incomplete Markets, Liquidation Risk and the Term Structure of Interest Rates," *Journal of Economic Theory*, 148(6), 2483–2519.

Challe, E., J. Matheron, X. Ragot, and J. F. Rubio-Ramirez (2017): "Precautionary Saving and Aggregate Demand," *Quantitative Economics*, 8(2), 435–478.

Challe, E., and X. Ragot (2016): “Precautionary Saving over the Business Cycle,” *Economic Journal*, 126(590), 135–164.

Chetty, R. (2009): “Sufficient Statistics for Welfare Analysis: A Bridge between Structural and Reduced-Form Methods,” *Annual Review of Economics*, 1(1), 451–488.

Chetty, R., and A. Finkelstein (2013): “Social Insurance: Connecting Theory to Data,” in *Handbook of Public Economics*, ed. by E. Saez, M. Feldstein, R. Chetty, and A. J. Auerbach, vol. 5, chap. 3, pp. 111–193.

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(3), 471–475.

Chien, Y., H. Cole, and H. Lustig (2011): “A Multiplier Approach to Understanding the Macro Implications of Household Finance,” *Review of Economic Studies*, 78(1), 199–234.

Cooley, T., and E. Prescott (1995): “Economic Growth and Business Cycles,” in *Frontiers of Business Cycle Research*, ed. by T. Cooley, chap. 1.

Dávila, J., J. H. Hong, P. Krusell, and J.-V. Ríos-Rull (2012): “Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks,” *Econometrica*, 80(6), 2431–2467.

Den Haan, W. J. (2010): “Comparison of Solutions to the Incomplete Markets Model with Aggregate Uncertainty,” *Journal of Economic Dynamics and Control*, 34(1), 4–27.

Dyrda, S., and M. Pedroni (2016): “Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks,” Working Paper, University of Toronto.

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

Heathcote, J. (2005): “Fiscal Policy with Heterogeneous Agents and Incomplete Markets,” *Review of Economic Studies*, 72(1), 161–188.

Hopenhayn, H. A., and E. C. Prescott (1992): “Stochastic Monotonicity and Stationary Distributions for Dynamic Economies,” *Econometrica*, 60(6), 1387–1406.

Huggett, M. (1993): “The Risk Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies,” *Journal of Economic Dynamics and Control*, 17(5-6), 953–969.

Imrohoroğlu, A. (1989): “Cost of Business Cycles with Indivisibilities and Liquidity Constraints,” *Journal of Political Economy*, 97(6), 1364–1383.

Kolsrud, J., C. Landais, P. Nilsson, and J. Spinnewijn (2018): “The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden,” *American Economic Review*, 108(4-5), 985–1033.

Krueger, D., and A. Ludwig (2016): “On the Optimal Provision of Social Insurance: Progressive Taxation Versus Education Subsidies in General Equilibrium,” *Journal of Monetary Economics*, 77, 72–98.

Krueger, D., K. Mittman, 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, pp. 143–184. Cambridge University Press.

Krusell, P., T. Mukoyama, and A. Sahin (2010): “Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations,” *Review of Economic Studies*, 77(4), 1477–1507.

Krusell, P., and A. A. J. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Landais, C., P. Michaillat, and E. Saez (2018): "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory," *American Economic Journal: Economic Policy*, 10(2), 152–181.

Maliar, L., S. Maliar, and F. Valli (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell-Smith Algorithm," *Journal of Economic Dynamics and Control*, 34(1), 42–49.

Marcet, A., and R. Marimon (2011): "Recursive Contracts," Working Paper, European University Institute.

McKay, A., and R. Reis (2016): "Optimal Automatic Stabilizers," NBER Working Paper 22359, National Bureau of Economic Research.

Mitman, K., and S. Rabinovich (2015): "Optimal Unemployment Insurance in an Equilibrium Business-Cycle Model," *Journal of Monetary Economics*, 71, 99–118.

Nuño, G., and B. Moll (2018): "Social Optima in Economies with Heterogeneous Agents," 48, 150–180.

Reiter, M. (2009): "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 33(3), 649–665.

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

Shimer, R. (2003): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95(1), 25–49.

Tao, T. (2011): *An Introduction to Measure Theory*. American Mathematical Society.

Winberry, T. (2018): "A Method for Solving and Estimating Heterogeneous Agent Macro Models," *Quantitative Economics*, 9(3).

Appendix

A Projection

Transition probabilities. We start by reformulating the transition probability $\Pi^S_{\tilde{h}h,t}$. Because of bin heterogeneity, agents with histories $s_1^{t-1} \neq s_2^{t-1}$ in $\tilde{h}$ can face different probabilities of transitioning to h. However, we can partition each $\tilde{h}$ into $N_{\tilde{h},t}$ elements denoted by $\{\tilde{h}_i\}_{i=1,\ldots,N_{\tilde{h},t}}$ such that for all $s^{t-1} \in \tilde{h}_i$, the probability of transitioning from s^{t-1} to a bin h is the same and denoted by $\Pi^S_{\tilde{h}_i h,t}$. Using the definition of conditional probability, we have: $\mu_t(s^t \in h | s^{t-1} \in \tilde{h}) = \frac{\mu_t(\{s^t \in h\} \cap \{s^{t-1} \in \tilde{h}\})}{\mu_t(s^{t-1} \in \tilde{h})} = \frac{1}{\mu_t(s^{t-1} \in \tilde{h})} \sum_{i=1}^{N_{\tilde{h},t}} \mu_t(\{s^t \in h\} \cap \{s^{t-1} \in \tilde{h}_i\})$. Using the definition of conditional probability again, and the fact that $\mu_t(s^t \in h | s^{t-1} \in \tilde{h}_i) = \Pi_{\tilde{h}_i h,t}$ is by construction independent of s^{t-1}, we get:

$$\Pi^S_{\tilde{h}h,t} = \sum_{i=1}^{N_{\tilde{h},t}} \Pi^S_{\tilde{h}_i h,t} \frac{S_{\tilde{h}_i,t-1}}{S_{\tilde{h},t-1}}. \tag{50}$$

In the case of explicit partitions, there is no such heterogeneity and $N_{\tilde{h},t} = 1$ for all $\tilde{h}$.

Bin sizes. For $h \in \mathcal{H}$, the bin size is: $S_{h,t} = \sum_{\tilde{h} \in \mathcal{H}} \sum_{s^{t-1} \in \tilde{h}, s^t \in h} \mu_t(s^t | s^{t-1}) \mu_{t-1}(s^{t-1})$. Using the partitions of $\tilde{h}$ as above, we obtain:

$$S_{h,t} = \sum_{\tilde{h} \in \mathcal{H}} \sum_{i=1}^{N_{\tilde{h},t}} \sum_{s^{t-1} \in \tilde{h}_i} \underbrace{\mu_t(s^t \in h | s^{t-1})}_{=\Pi^S_{\tilde{h}_i h,t}} \mu_{t-1}(s^{t-1}) = \sum_{\tilde{h} \in \mathcal{H}} \sum_{i=1}^{N_{\tilde{h},t}} \Pi^S_{\tilde{h}_i h,t-1} \underbrace{\sum_{s^{t-1} \in \tilde{h}_i} \mu_{t-1}(s^{t-1})}_{=S_{\tilde{h}_i,t-1}}.$$

Using (50), we finally deduce: $S_{h,t} = \sum_{\tilde{h} \in \mathcal{H}} \Pi^S_{\tilde{h}h,t} S_{\tilde{h},t-1}$. For an explicit partition, this relationship is unchanged (but $N_{\tilde{h},t} = 1$ in the proof).

Projecting lagged variables. We have $\mathbb{A}_{h,t}[X_{t-1}] = \frac{1}{S_{h,t}} \sum_{\tilde{h} \in \mathcal{H}, s^{t-1} \in \tilde{h}} X_{t-1}\left(s^{t-1}\right) \mu_t(s^t \in h | s^{t-1}) \mu_{t-1}(s^{t-1})$. Using a partition of $\tilde{h}$, and after some algebraic operations, we obtain:

$$\mathbb{A}_{h,t}[X_{t-1}] = \frac{1}{S_{h,t}} \sum_{\tilde{h} \in \mathcal{H}} \underbrace{\frac{\sum_{i=1}^{N_{\tilde{h},t}} \Pi^S_{\tilde{h}_i h} S_{\tilde{h}_i,t-1} X_{\tilde{h}_i,t-1}}{S_{\tilde{h},t-1} X_{\tilde{h},t-1}}}_{=\Pi^X_{\tilde{h}h,t}} S_{\tilde{h},t-1} X_{\tilde{h},t-1}, \tag{51}$$

Note that $\Pi^X_{\tilde{h}h,t} \in [0,1]$ and $\sum_{h \in \mathcal{H}} \Pi^X_{\tilde{h}h,t} = 1$ and that $\left(\Pi^X_{\tilde{h}h,t}\right)_{\tilde{h}h}$ defines a transition matrix.

For an explicit partition, $N_{\tilde{h},t} = 1$ and we can simplify equation (51) with $\Pi^X_{\tilde{h}h,t} = \Pi^S_{\tilde{h}h,t}$.

Projecting a conditional expectation. Using the same technique as for $S_{h,t}$, we obtain:

$$\mathbb{A}_{h,t}[\mathbb{E}_t X_{t+1}\left(s^{t+1}\right)] \quad = \quad \mathbb{E}_t \sum_{\tilde{h} \in \mathcal{H}} \underbrace{\Pi^S_{h\tilde{h},t+1} \frac{\sum_{s^t \in h} X_{t+1}\left((s^t, s_{\tilde{h}})\right) \frac{\mu_t(s^t)}{S_{h,t}}}{X_{t+1,\tilde{h}}}}_{=\Pi^u_{\tilde{h}h,t}} X_{t+1,\tilde{h}}. \tag{52}$$

For explicit partitions, the expression barely simplifies but in the absence of bin heterogeneity (i.e., $\forall s^t \in h,\ X_{t+1}\left((s^t, s_{\tilde{h}})\right) = X_{t+1,\tilde{h}}$), we have $\mathbb{A}_{h,t}[\mathbb{E}_t X_{t+1}\left(s^{t+1}\right)] = \mathbb{E}_t \sum_{\tilde{h} \in \mathcal{H}} \Pi^S_{h\tilde{h},t+1} X_{t+1,\tilde{h}}$.

B Proof of Proposition 2

We denote by $s^\infty = (s_0, s_1, \ldots)$ an infinite idiosyncratic history. The history s^∞ can be seen as a Markov chain where each s_i takes a value in $(\mathcal{S}, \mathcal{F}_\mathcal{S})$, where $\mathcal{F}_\mathcal{S}$ are the σ-algebras generated by $\mathcal{S}$ (which is finite in our case). The whole chain s^∞ lies in the sequence space $\Omega = \mathcal{S} \times \mathcal{S} \times \ldots = \mathcal{S}^\infty$ endowed with the product σ-algebra $\mathcal{F}_\infty = \mathcal{F}_\mathcal{S}^\infty$ and the measure μ_∞. We recall that the set Ω is uncountably infinite and has the cardinality of the continuum. Such a measure can be seen as the infinite product measure that coincides with the standard Markov distribution for any finite sequence. This measure exists even when the state-space $\mathcal{S}$ is infinite (and even when the support of the time index is a subset of $\mathbb{R}$) and can be shown to be the limit of the product measure of transition kernels (generalizing the transition matrix in the case of a non-finite state-space $\mathcal{S}$). An important feature of the infinite product measure is that it is consistent with the usual Markov measure for any finite sequence. The proof of the existence of the infinite product measure is in general not trivial and relies on the Kolmogorov extension theorem (see Tao 2011, Theorem 2.4.3). In our case of a finite Markov chain, the infinite measure is uniquely determined by its initial distribution and its transition matrix (see Brémaud 2014, Theorem 1.1).

We now consider the probability space $(\Omega, \mathcal{F}, \mu_\infty)$. Let $\mathcal{H}_n$ be a partition of idiosyncratic histories. A history-bin $h_n \in \mathcal{H}_n$ can then be seen as a subset of Ω and the projection of the variable X in equation (20) can be written as: $X_{h_n,t} = \int_{s^\infty \in h_n} X_t(s^\infty) \frac{\mu_\infty(ds^\infty)}{\int_{s^\infty \in h_n} \mu_\infty(ds^\infty)}$.

We consider the filtration associated with the partition $\mathcal{H}_n$ that we denote by $\mathcal{F}_n$. The conditional probability $\mathbb{E}\left[X_t|\mathcal{F}_n\right]$ is a random variable, such that for any event $s^\infty \in \Omega$, we have:

$$\mathbb{E}\left[X_t|\mathcal{F}_n\right]_{s^\infty} = X_{h_n,t}, \tag{53}$$

where h_n is the unique partition element containing s^∞. In other words, the restriction of the conditional expectation to h_n coincides with $X_{h_n,t}$. We have two additional properties on the filtration sequence $(\mathcal{F}_n)_{n\geq 0}$.

1. The increasing partition sequence $(\mathcal{H}_n)_{n\geq 0}$ is such that $\mathcal{H}_{n+1}$ is a refinement of $\mathcal{H}_n$, in the sense that any element of $\mathcal{H}_n$ is a union of elements of $\mathcal{H}_{n+1}$. The filtration sequence $(\mathcal{F}_n)_{n\geq 0}$ is thus increasing: $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.

2. The partition sequence $(\mathcal{H}_n)_{n\geq 0}$ converges to the atoms of Ω, which implies $(\mathcal{F}_n) \uparrow \mathcal{F}_\infty$.

To conclude the convergence proof, we apply the convergence theorem for conditional expectation (see Theorem 11.2 in Billingsley 1965), which yields: $\mathbb{E}\left[X|\mathcal{F}_n\right]_{s^\infty} \to X(s^\infty)$.

C Proof of Corollary 1

Let $(\mathcal{H}_n)_{n\geq 0}$ be an increasing partition sequence. First, since $\xi^f_{h_n,t} = \frac{\mathbb{A}_{h_n,t}[f(X_t)]}{f(X_{h_n,t})}$, we have using definition (53), for any $h_n \in \mathcal{H}_n$, $\xi^f_{h_n,t} = \frac{\mathbb{E}[f(X)|\mathcal{F}_n]_{s^\infty}}{f(\mathbb{E}[X|\mathcal{F}_n]_{s^\infty})}$, where again h_n is the unique partition element containing s^∞. Using the convergence result we have just proved and the continuity of f, we have $\xi^f_{h_n,t} \to \frac{f(X(s^\infty))}{f(X(s^\infty))} = 1$.

Regarding the probability $\Pi^X_{\tilde{h}_n h_n,t}$ defined in (51), we have, for any $h_n, \tilde{h}_n \in \mathcal{H}$:

$$\Pi^X_{\tilde{h}_n h_n,t} - \Pi^S_{\tilde{h}_n h_n,t} = \frac{\sum_{i=1}^{N_{\tilde{h}_n,t}} \Pi^S_{\tilde{h}_{n,i} h_n} S_{\tilde{h}_{n,i},t-1}(X_{\tilde{h}_{n,i}t-1} - X_{\tilde{h}_n,t-1})}{S_{\tilde{h}_n,t-1} X_{\tilde{h}_n,t-1}},$$

from which we deduce that: $|\Pi^X_{\tilde{h}_n h_n,t} - \Pi^S_{\tilde{h}_n h_n,t}| \leq K \sum_{i=1}^{N_{\tilde{h}_n,t}} |X_{\tilde{h}_{n,i},t-1} - X_{\tilde{h}_n,t-1}|$ for some $K > 0$. Using definition (53), we obtain that $|\Pi^X_{\tilde{h}_n h_n,t} - \Pi^S_{\tilde{h}_n h_n,t}| \leq K \sum_{i=1}^{N_{\tilde{h}_n,t}} |\mathbb{E}\left[X_t|\mathcal{F}_{n,i}\right]_{s^\infty} - \mathbb{E}\left[X_t|\mathcal{F}_n\right]_{s^\infty}|$ (where $\mathcal{F}_{n,i}$ is the filtration corresponding to the sub-partition $(h_{n,i})_{i,n}$). The convergence result implies that the absolute difference of expectation converges to 0 (both expectations converge to the same value) and $|\Pi^X_{\tilde{h}_n h_n,t} - \Pi^S_{\tilde{h}_n h_n,t}| \to 0$.

Finally, we have for $\Pi^u_{\tilde{h}h,t}$ defined in (52):

$$\Pi^u_{h_n\tilde{h}_n,t+1} - \Pi^S_{h_n\tilde{h}_n,t+1} = \frac{\Pi^S_{h_n\tilde{h}_n,t+1}}{X_{t+1,\tilde{h}_n}}\left(\frac{1}{S_{h_n,t}} \sum_{s^t \in h_n} X_{t+1}(s^t, s_{\tilde{h}_n})\mu_t(s^t) - X_{t+1,\tilde{h}_n}\right),$$

or using definition (53), $\Pi^u_{h_n\tilde{h}_n,t+1} - \Pi^S_{h_n\tilde{h}_n,t+1} = \frac{\Pi^S_{h_n\tilde{h}_n,t+1}}{X_{t+1,\tilde{h}_n}}(\mathbb{E}[X_{t+1}|\mathcal{F}_{n,t}]_{s^\infty} - \mathbb{E}[X_{t+1}|\mathcal{F}_n]_{s^\infty})$ – where we add to the subscript t to the filtration to distinguish the conditional expectation. The convergence result implies then that $\Pi^u_{h_n\tilde{h}_n,t+1} - \Pi^S_{h_n\tilde{h}_n,t+1} \to 0$.

D Summary of the dynamics of the optimal allocation

The dynamics system characterizing the optimal allocation can be written as follows:

$$\begin{aligned}
h \in \mathcal{H} &: l_{h,t} = \chi^\varphi (1-\tau_t)^\varphi w_t^\varphi y_h^\varphi 1_{e_h=e}, \\
h \in \mathcal{H} &: c_{h,t} + a_{h,t} \leq (1+r_t)\tilde{a}_{h,t} + ((1-\tau_t)1_{e_h=e} + \phi_t 1_{e_h=u})\, l_{h,t} y_h w_t, \\
h \in \mathcal{H} &: \tilde{a}_{h,t} = \sum_{\tilde{h}\in\mathcal{H}} \tilde{\Pi}^a_{\tilde{h}h} \Pi_{y_{\tilde{h}} y_h} \Pi_{y_h y_{\tilde{h}}} \Pi_{e_h e_{\tilde{h}},t} \frac{S_{\tilde{h}}}{S_h} a_{\tilde{h},t}, \\
h \in \mathcal{H} &: \Psi_{h,t} = \xi^U_h U_{c,h,t} - (\lambda_{h,t} - (1+r_t)\Lambda_{h,t})\, \xi^u_h U_{cc,h,t}, \\
h \notin \mathcal{C} &: \xi^u_h U_c(c_{h,t}, l_{h,t}) = \beta(1+r) \sum_{\tilde{h}\in\mathcal{H}} \tilde{\Pi}^u_{h\tilde{h}} \Pi_{y_h y_{\tilde{h}}} \Pi_{e_h e_{\tilde{h}},t} \xi^u_{\tilde{h}} U_c(c_{\tilde{h},t}, l_{\tilde{h},t}),
\end{aligned} \tag{54}$$

$$h \notin \mathcal{C} : \Psi_{h,t} = \beta \sum_{\tilde{h},} \tilde{\Pi}^a_{h\tilde{h}} \Pi_{y_h y_{\tilde{h}}} \Pi_{e_h e_{\tilde{h}},t} \mathbb{E}_t \left[(1+r_{t+1})\Psi_{\tilde{h},t+1}\right] + \beta \frac{1-\alpha}{\varphi} \mathbb{E}_t \left[\frac{1}{L_{t+1}} \left(\frac{K_t}{L_{t+1}}\right)^{\alpha-1} \right.$$

$$\left. \times \sum_{\tilde{h}} \left((1-\tau_{t+1}) 1_{e_{\tilde{h}}=e} + \frac{S_{e,t+1}}{S_{u,t+1}} \left(\tau_{t+1}\left(1+\varphi\right) - 1\right) 1_{e_{\tilde{h}}=u} \right) \Psi_{\tilde{h},t+1} S_{\tilde{h},t+1} l_{\tilde{h},t+1} y_{\tilde{h}} \right],$$

$$h \in \mathcal{C} : a_{h,t} = -\bar{a} \text{ and } \lambda_{h,t} = 0,$$

$$K_t = \sum_{h\in\mathcal{H}} S_{h,t} a_{h,t}, \text{ and } L_t = \sum_{h\in\mathcal{H}} S_{h,t} y_h l_{h,t},$$

$$\sum_{h\in\mathcal{H}} S_{h,t} c_{h,t} + K_t = Y_t + K_{t-1} \text{ and } \phi_t \sum_{h\in\mathcal{E}_t} S_{h,t} y_h l_{h,e,t} = \tau_t \sum_{h\in\mathcal{H}} S_{h,t} y_h l_{h,t},$$

$$r_t = \alpha Z_t \left(\frac{K_{t-1}}{L_t}\right)^{\alpha-1} - \delta \text{ and } w_t = (1-\alpha) Z_t \left(\frac{K_{t-1}}{L_t}\right)^{\alpha},$$

$$\sum_{h\in\mathcal{H}} S_{h,t} \Lambda_{h,t} \xi^u_h U_{c,h,t} = -\sum_{h\in\mathcal{H}} S_{h,t} \Psi_{h,t} \tilde{a}_{h,t} + \frac{1}{\alpha\varphi} \frac{K_{t-1}}{L_t}$$

$$\times \sum_{h\in\mathcal{H}} \left(-(1-\tau_t) 1_{e_h=e} + \frac{S_{e,t}}{S_{u,t}} \left(1 + \alpha\varphi - (1+\varphi)\tau_t\right) 1_{e_h=u} \right) S_{h,t} \Psi_{h,t} l_{h,t} \tilde{y}. \tag{55}$$

Two remarks are in order. First, for sake of simplicity, in equation (54) we define the beginning-of-period wealth in any bucket h as $\tilde{a}_{h,t}$. Second, the optimal replacement rate and labor tax rate is given by equation (55), corresponding to the first-order condition of the program (42).

E Matrix representation at the steady-state

Before turning to the matrix representation, we introduce the following notation:

$\circ$ is the Hadamard product, $\otimes$ is the Kronecker product, $\times$ is the usual matrix product.

For any vector V, we denote by $diag\,(V)$ the diagonal matrix with V on the diagonal. We assume that there are $N^{tot} = N \times Card(\mathcal{Y}) \times 2$ elements in $\mathcal{H}$, which can be identified by the $(b, y, e)_{b=1,\ldots,N;y=1,\ldots,Card(Y);e=1,2}$, where $e = 1$ if the agent is unemployed and $e = 2$ if she is employed, and N is the number of wealth bins. Equivalently, elements of $\mathcal{H}$ can be identified by $h = b + N \times (y-1) + (e-1) \times Card(\mathcal{Y}) \times N$, such that $h = 1, \ldots, N^{tot}$. In other words, we stack the wealth indices, then the productivity indices, and then the employment indices.

We derive Lagrange multipliers as a function of the the steady-state solution (i.e., allocations and prices), which is assumed to be known. Let $\mathbf{S}$ be the N^{tot}-vector of steady-state bin sizes. Similarly, let $\mathbf{a}$, $\mathbf{c}$, $\boldsymbol{\ell}$, $\boldsymbol{\nu}$, $\mathbf{U_c}$, $\mathbf{U_{cc}}$, $\boldsymbol{\xi^U}$, and $\boldsymbol{\xi^u}$ be the N^{tot}-vectors of end-of-period wealth, consumption, labor supply, Lagrange multipliers, marginal utilities, derivative of the marginal utility, and correcting parameters, respectively. Also let:

$$\mathbf{W} = w \begin{bmatrix} \phi \\ 1-\tau \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ \vdots \\ y_{n_y} \end{bmatrix} \otimes 1_B,\ \mathbf{L}_e = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ \vdots \\ y_{n_y} \end{bmatrix} \otimes 1_B,\ \mathbf{L}_u = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ \vdots \\ y_{n_y} \end{bmatrix} \otimes 1_B,$$

In addition, define as $\mathbb{P}$ the diagonal matrix having 1 on the diagonal at row h if and only if the "agent" h is not credit constrained (i.e., $\nu_h = 0$), and 0 otherwise. Similarly, define $\mathbb{P}^c = \mathbf{I} - \mathbb{P}$, where $\mathbf{I}$ is the $(N^{tot} \times N^{tot})$-identity matrix.

From transitions across elements of the partition we can compute the $(N^{tot} \times N^{tot})$-matrices Π^S, Π^a, Π^u such that the market economy is summarized by:

$$\begin{aligned} \mathbf{S} &= \Pi^S \mathbf{S}, \\ \mathbf{S} \circ \mathbf{c} + \mathbf{S} \circ \mathbf{a} &= (1+r)\Pi^a (\mathbf{S} \circ \mathbf{a}) + (\mathbf{S} \circ \mathbf{W} \circ \boldsymbol{\ell}), \\ \mathbb{P}\boldsymbol{\xi}^u \circ u'(\mathbf{c}) &= \mathbb{P}\beta(1+r)\Pi^u (\boldsymbol{\xi}^u \circ u'(\mathbf{c})) + \boldsymbol{\nu}, \\ \mathbb{P}^c \mathbf{a} &= -\bar{a} 1_{N^{tot}\times 1}, \\ \left(\frac{r+\delta}{\alpha}\right)^{\frac{1}{\alpha-1}} \mathbf{L_e}^\top \times \mathbf{S} &= \mathbf{S}^\top \times \mathbf{a}, \\ \tau = \phi \frac{\mathbf{L_u}^\top \times \mathbf{S}}{\mathbf{L_e}^\top \times \mathbf{S}}, &\text{ and } w = (1-\alpha)\left(\frac{r+\delta}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}. \end{aligned} \tag{56}$$

If we denote by $\boldsymbol{\lambda}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ the vectors associated with the Lagrange multipliers, we have: $\boldsymbol{\Psi} = \boldsymbol{\xi}^{\mathbf{U}} \circ \mathbf{U_c} - (\boldsymbol{\lambda} - (1+r)\boldsymbol{\Lambda}) \circ \boldsymbol{\xi}^{\boldsymbol{u}} \circ \mathbf{U_{cc}}$. Define the matrix Π^Λ by $\Pi^\Lambda_{\tilde{k}k} = \frac{S_k \Pi^u_{k\tilde{k}}}{S_{\tilde{k}}}$, such that $\boldsymbol{\Lambda} = \Pi^\Lambda \boldsymbol{\lambda}$, and the matrix Π^Ψ by:

$$\begin{aligned} \Pi^\Psi_{k\tilde{k}} &= \beta(1+F_K)\Pi^a_{k\tilde{k}} \\ &+ \beta \frac{1-\alpha}{\alpha\varphi} \frac{1}{L} (r+\delta) \left((1-\tau) 1_{\tilde{k} > \frac{N^{tot}}{2}} + \frac{S_e}{S_u} (\tau(1+\varphi) - 1) 1_{\tilde{k} \leq \frac{N^{tot}}{2}} \right) S_{\tilde{k}} l_{\tilde{k}} (\mathbf{L}^e + \mathbf{L}^u)_{\tilde{k}}. \end{aligned}$$

Note that $1_{\tilde{k} > \frac{N^{tot}}{2}}$ represents employed agents and $1_{\tilde{k} \leq \frac{N^{tot}}{2}}$ unemployed agents. One can check that $\mathbb{P}\boldsymbol{\Psi} = \mathbb{P}\Pi^\Psi \boldsymbol{\Psi}$. The key point is that Π^Ψ and Π^Λ are known from the allocations.

After simple matrix algebra, we find the value of the vector of the Lagrange multiplier $\boldsymbol{\lambda}$ as:

$$\boldsymbol{\lambda} = \left[\mathbb{P}^c + \mathbb{P}\left(\mathbf{I} - \Pi^\Psi\right)\left(diag\left(\boldsymbol{\xi}^{\boldsymbol{u}} \circ \mathbf{U_{cc}}\right) \times \left(\mathbf{I} - (1+F_K)\Pi^\Lambda\right)\right)\right]^{-1} \mathbb{P}\left(\mathbf{I} - \Pi^\Psi\right)\left(\boldsymbol{\xi}^{\mathbf{U}} \circ \mathbf{U_c}\right). \tag{57}$$

Importantly, the whole right-hand side can be deduced from the Bewley allocation. This makes the computation of λ straightforward. We then deduce $(\mathbf{\Lambda}, \mathbf{\Psi})$ with:

$$\mathbf{\Lambda} = \Pi^{\Lambda}\lambda, \text{ and } \mathbf{\Psi} = \xi^{U} \circ \mathbf{U}_c - \xi^{\mathbf{u}} \circ \mathbf{U}_{\mathbf{cc}} \circ \left(\mathbf{I} - (1+F_K)\Pi^{\Lambda}\right)\lambda. \tag{58}$$

Using equations (57)–(58), we can derive Lagrange multipliers from the real allocation only.

The condition (55) for the optimality of the real allocation can be simplified into:

$$\mathbf{V}^{\top} \times 1_{N^{tot}\times 1} = 0, \tag{59}$$

$$\text{where: } \mathbf{V} \equiv \frac{\varphi\alpha}{K}(\mathbf{S}\circ\Psi\circ\tilde{\mathbf{a}} + \mathbf{S}\circ\Lambda\circ\xi^{\mathbf{u}}\circ\mathbf{U}_c) + \frac{1-\tau}{L}\mathbf{S}\circ\Psi\circ\mathbf{L}^e - (1-\tau+\varphi(\alpha-\tau))\frac{S_e}{S_u}\mathbf{S}\circ\Psi\circ\mathbf{L}^u.$$

F Algorithm for solving the Ramsey problem

The algorithm for computing Ramsey policies is provided for implicit partitions, but can be easily modified for explicit partitions.

1. As explained in Section 3, we consider a partition $\mathcal{H}$ of idiosyncratic histories. We use the same wealth partition $\mathcal{B}$ for all idiosyncratic states. Assuming that the wealth partition $\mathcal{B}$ has N elements, the partition $\mathcal{H}$ then counts $Card(\mathcal{Y}) \times 2 \times N$ different history bins.

2. We set a reasonable initial value for the replacement rate ϕ.

   (a) We solve the general Bewley model for the value of ϕ.

   (b) We use the steady-state results to compute the model projection on the partition $\mathcal{H}$. In particular we compute the constant quantities $\left(\xi_h^u, \xi_h^U, \widetilde{\Pi}^u_{h\tilde{h}}, \widetilde{\Pi}^a_{\tilde{h}h}, \widetilde{\Pi}_{\tilde{h}h}\right)_{\tilde{h},h\in\mathcal{H}}$.

   (c) We determine the steady-state values of the multipliers $(\lambda_h)_{h\in\mathcal{H}}$ and $(\Lambda_h)_{h\in\mathcal{H}}$, and of the social value of liquidity $(\Psi_h)_{h\in\mathcal{H}}$ using equations (57) and (58).

3. We iterate on ϕ and repeat Step 2 until equality (59) holds.

Once the steady-state and the partition have been determined, it is easy to simulate the model using standard perturbation techniques with existing software such as Dynare (see Adjemian et al., 2011). The simulation of the whole optimal allocation for the calibrated economy with aggregate shocks requires less than 1 minute on a standard laptop.