The Welfare of Nations: Social Preferences and the Macroeconomy*

François Le Grand   Xavier Ragot   Diego Rodrigues†

November 2024

We present an aggregation theory for the Social Welfare Function (SWF) that enables its empirical estimation. Agents have heterogeneous perceptions of how the planner should value the welfare of other agents, which results in so-called Individual Welfare Functions (IWFs), shaped by their life experiences. The SWF is constructed as an aggregation of IWFs, weighted by the political weights of each group of agents. We then develop an estimation strategy to identify the SWF and IWFs based on the observed levels of capital tax, consumption tax, labor tax, and public debt. This strategy extends the inverse optimal approach to a general equilibrium heterogeneous-agent model. The application of our methodology to France and the United States shows that France's SWF is more Egalitarian and places greater emphasis on low-income individuals, contrasting with the United States' SWF, which is more Libertarian and assigns greater weight to high-income individuals. Finally, we simulate the fiscal system of the United States under the assumption of adopting the French SWF. Our findings indicate that the SWF significantly shapes the observed fiscal system and equilibrium inequality.

Keywords : Social Welfare Function, Inequality, Fiscal systems.

JEL codes : E61, E62, E32. papers

*We thank Roland Bénabou, Anmol Bhandari, Christian Bayer, Julia Cagé, Edouard Challe, Eduardo Davilá, Aurélien Eyquem, Antoine Ferey, Raquel Fernandez, Jonathan Heathcote, Dirk Krueger, Alexander Monge-Naranjo, Giunluca Violante, Nicolas Werquin, and participants at EEA, NBER SI (IM), CASFI conference in Bonn, T2M conference. We also acknowledge financial support from ANR (ANR-20-CE26-0018 IRMAC).

†LeGrand: Rennes School of Business; `francois.le-grand@rennes-sb.com`. Ragot: SciencesPo-CNRS, OFCE, and CEPR; `xavier.ragot@gmail.com`. Rodrigues: SciencesPo; `diego.desousarodrigues@sciencespo.fr`.

1 Introduction

Countries differ widely in terms of their fiscal system and inequalities. For instance, the average mandatory levies in France were 43% of its GDP from 1995 to 2007, while they were around 26% in the US for the same period. Various explanations can rationalize these international differences. First, production technologies can be different across countries, and so do the distortions generated by taxation. Second, individual preferences over consumption and leisure, as well as social preferences regarding redistribution, can differ across countries. Third, even if technology and preferences are the same, the political system that selects and implements actual policies can differ across countries, generating outcomes that are observationally equivalent to alternative social preferences. The goal of this paper is to disentangle these three explanations and to identify the contribution of social preferences in the design of actual tax systems.

To do so, we present a theory of aggregation of individual preferences into social preferences, which can then be estimated using available data. This Bewley-type aggregation theory is based on four assumptions. First, individuals have their own view of how the planner should care about the welfare of all people in the society. This corresponds to the so-called ethical preferences (Arrow, 1951, Harsanyi, 1955, or Sen, 1977). To quote Harsanyi (1955), ethical preferences are indeed defined as an agent choosing "what he prefers only in those possibly rare moments when he forces a special impartial and impersonal attitude on himself."[1] These ethical preferences are represented by Individual Welfare Functions (IWFs), which are agent-specific. IWFs are thus heterogeneous, consistently with empirical investigations (Gaertner and Schokkaert, 2012, Fehr et al., 2013 or Stantcheva, 2021). Our second assumption is about the source of heterogeneity for IWFs: The IWFs are the outcome of the life experience of each agent. Following the Bewley tradition, we assume that agents' relevant life experience is their economic history. Our construction could easily manage other dimensions of heterogeneity, but this representation is already sufficiently rich for us to discuss a wide variety of political implications of our estimation. Third, we assume that the ethical preferences are of the weighted utilitarism type. Agents value the utility of others agents by attributing some weights to the individual utility of the latter. This representation is known to be flexible enough to embed moral and political concerns (including Libertarian, Egalitarian or Utilitarian ones) into the planner's motives (Saez and Stantcheva, 2016 among others). Our fourth assumption is that the Social Welfare Function (SWF) that the planner uses to design the fiscal policies is the (possibly biased) aggregation of the heterogeneous IWFs. Indeed, agents can have heterogeneous political power in their ability to affect the planners' decisions. Consistently with the literature on political inequality (e.g., Cage, 2024 for a discussion), this could for instance be explained by lobbying powers, participation rates in elections, or the representation of voters and their representatives. This construction yields a SWF that is stationary at the steady state, although agents' ethical preferences can change with their own history. The SWF can thus be estimated in the data, and can also be related to the concepts of public finance literature such as the Social Marginal Welfare Weights (SMWW) and the Marginal Value of Public Funds (MVPF). As a final remark, we do not require the SWF to fulfill the Pareto principle. Since Sen (1970) or Kaplow and Shavell (2001), it is

[1]These ethical preferences have a long tradition. They are the preferences of the "impartial spectator" of Adam Smith (1759) or preferences under the "Veil of ignorance" of Rawls (1971).

indeed known that allowing the planner's objective to account for moral or political concerns can imply SWFs, which do not fulfill the Pareto principle. We first present a very simple model, where this construction is transparent.

We then develop a quantitative methodology to estimate SWFs and IWFs from the data. We apply it to France and the US, which have the advantage of being very different in terms of taxation and inequality. This methodology extends the standard inverse optimal approach (see the literature review below) to a general equilibrium heterogeneous-agent model (à la Bewley-Huggett-Aiyagari). We start with showing that a fiscal system, composed of a progressive labor tax, a capital tax, a consumption tax and public debt, combined with an empirically relevant income risk, can closely reproduce income and wealth inequality in 2007, both in France and in the US. We choose the period 2007 to exclude the financial crisis and the subsequent Covid-19 crisis, which led to significant transitory changes in fiscal structures. We then consider an heterogeneous-agent model, where agents face income risk and where the previous fiscal system finances a public good. A Ramsey planner sets the fiscal policy to maximize a SWF, which results from our construction. To solve this intertemporal program in general equilibrium, we use the truncation method, which has been progressively developed in LeGrand and Ragot (2022a, 2023). We here extend this approach to be able to consider general disutility of labor, instead of the GHH case of LeGrand and Ragot (2024). We finally use the first-order conditions (FOCs) of this Ramsey program to compute the SWF and the IWFs for which the observed allocation in France and in the US is a Ramsey steady state, which extends the standard inverse optimal approach to an intertemporal general equilibrium approach. The derivation of the Ramsey FOCs offers as a side benefit an intertemporal and general equilibrium expression of concepts of SMWW and MVPF.

The estimation of SWFs generates three sets of results. First, the SWFs in France and in the US are very different from each other. The US SWF weights increase with income and put the largest weight on high-income agents. The weights for middle-class agents are lower, and low-income agents have the lowest weight. This shape of the SWF weights is actually consistent with the decreasing shape of SMWW estimated in the literature, as the latter include marginal utilities, as we explain. Oppositely, the French SWF assigns the highest weight to low-income agents. However, the SWF weights are U-shaped and the weights first decrease with income, such that the middle-class has the lowest weight, while they then increase for high-income agents. Maybe not surprisingly, the French SWF is egalitarian at the bottom of the distribution, but also puts a high weight on very productive agents. Second, to understand the role of the SWF in shaping inequality, we simulate the optimal US fiscal system if the US were to adopt the French SWF, keeping individual preferences constant. We find that the Gini coefficient of wealth would decrease from 78% to 63%, approaching the French Gini coefficient of wealth of 68%. Consequently, social preferences are a primary driver of the fiscal system and household inequality. Third, we decentralize the aggregate SWF in heterogeneous IWFs to assess the heterogeneity of social preferences in each country. To do so, we first use turnout data in the US and in France to estimate political weights, following the political economy literature surveyed by Cage (2024). We then estimate the set of IWFs, which are the closest to the self-interested one and which are consistent with the SWF. We find that the middle class is mostly libertarian in the US and egalitarian in France, and that there is a substantial heterogeneity in IWFs within countries.

Related literature. Our paper is related to three streams of the literature: heterogeneous-agent macroeconomics, public finance, and social choice.

First, this paper contributes to the recent literature on optimal policies in heterogeneous-agent models. Early contributions, such as Aiyagari (1995) analyze general properties about capital tax in heterogeneous agent models. Aiyagari and McGrattan (1998) compute the optimal steady-state level of public debt. Dávila et al. (2012) show that the steady-state capital stock can be too low, solving for a constrained-efficient allocation. Some papers compute the optimal path of relevant instruments (Conesa et al., 2009 or Dyrda and Pedroni, 2022 more recently). Some papers rely on the FOCs of the Ramsey problem to solve for optimal policies (Bhandari et al., 2021; LeGrand and Ragot, 2022a; Açikgöz et al., 2022). We develop the algebra to connect the FOCs with public finance concepts and we use the recent developments of the truncation method of LeGrand and Ragot (2024) to apply the inverse optimal approach.

Second, the paper also contributes to the public finance literature, as it makes explicit the general equilibrium effects in the MSWW and in the MVPFs (Hendren and Sprung-Keyser, 2020 or Ferey et al., 2024 for a recent contribution). The inverse optimal approach, that we apply to an Ramsey program of an heterogeneous-agent model, is a common tool in public finance (Bargain and Keane, 2010; Bourguignon and Amadeo, 2015; Lockwood and Weinzierl, 2016; Hendren, 2020). Chang et al. (2018) also consider an heterogeneous-agent model to estimate inequality aversion across countries, but avoids the computation of a Ramsey program. Heathcote and Tsujiyama (2021) also estimate the SWF in a static environment, but allow for partial private insurance. In this literature, our contribution is implement inverse optimal approach in general equilibrium, contributing to fill the gap between macroeconomics and public finance.[2]

Third and finally, the paper contributes to the literature about SWF in heterogeneous-agent models. There is a vast theoretical literature in the social choice literature about the possible axiomatizations of SWFs – and the discussion of their motivations and their implications – that dates back to Arrow (1951), Harsanyi (1955), Sen (1970), Sen (1977) among many others. An empirical literature identifies possible relevant restrictions for SWFs from experiments (Gaertner and Schokkaert 2012 or Fehr et al., 2013) or from surveys (Stantcheva, 2021). Informed by this literature, we propose a construction of the SWFs as an aggregation of IWFs, which is flexible enough to allow for estimation. We consider our construction as a possible micro-foundation of the weights used to assess optimal policies, such as the Generalized Social Marginal Welfare Weights introduced by Saez and Stantcheva (2016).

The paper proceeds as follows. In Section 2, we present motivating evidence on the French and US fiscal system. We provide the basics of our Bewley construction of SWFs in Section 3 in the context of a simple model. The construction is generalized to an infinite horizon model in Section 4. Section 5 presents the environment in which we will compute the Ramsey program and conduct our estimation of SWF weights. Finally, the quantitative investigation is presented

[2]Papers in the experimental literature elicis social preferences using the spectator game (where a spectator is asked to split to resources between two unknown agents). For instance, Almås et al. (2020) find that US players are more Libertarian than Norwegians, who are more Egalitarian. These findings are broadly consistent with our analysis, which relies on an alternative identification strategy, based on the actual fiscal systems.

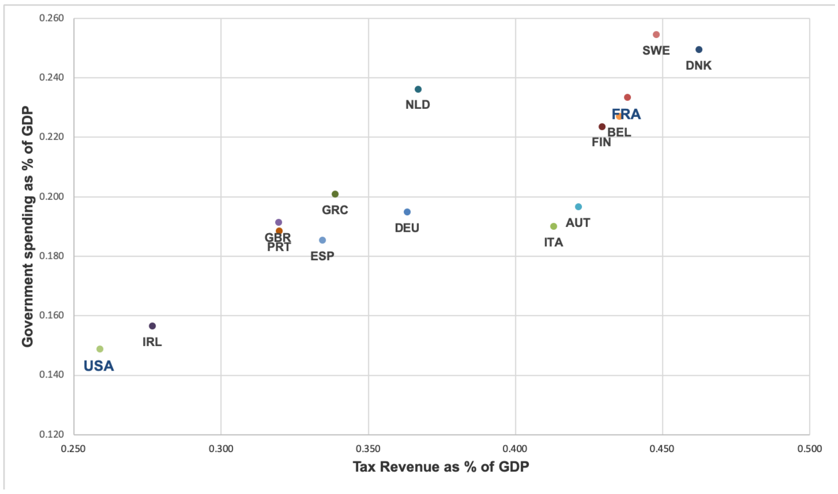


Figure 1: Government spending and tax revenues average from 1995 to 2021 (as a share of GDP). Source: own calculations.

in Section 6, while Section 7 concludes.

2 The fiscal structure in France and in the US

We report key statistics about the French and US fiscal systems. These two countries have the particularity of being among the most different OECD countries in terms of total taxation. On the one hand, France features one of the highest mandatory levies, while the US one of the lowest levies. This is confirmed in Figure 1, which reports government spending on final goods and services and tax revenues, both as a share of GDP, for the two countries. France and the US drastically differ with respect to the size of their governments: Both government spending and tax revenues are significantly higher in France than in the US. Moreover, the gap between tax revenues and government spending is larger in France than in the US. This reflects that the within-country redistribution – measured as the difference between tax revenues and government spending – is of larger magnitude in France than in the US.

We now turn to the details of the taxation system within each country. We focus on the average tax system from 1995 to 2007, before the 2008-crisis and the Covid crisis, which both were the sources of (so-far) transitory changes in fiscal systems.[3] We will use these values as a benchmark for calibrating our stationary equilibrium in the quantitative exercise of Section 6. For France and the US, we use the results of Trabandt and Uhlig (2011), who provide estimates for the period 1995-2007. Results for France and the US are gathered in Table 1, which also includes some elements related to inequalities.

[3]Actually, considering the period from 1995 to 2021, as shown in Figure 1, do not change the tax results significantly. However, for the sake of consistency, we chose to consider a period before macroeconomic shocks.

	Total taxes (%GDP)	$\tau^K(\%)$	$\tau^L(\%)$	$\tau^c(\%)$	B (%GDP)	G (%GDP)	Gini before redist.	Gini after redist.	Gini wealth
France	40	35	46	18	60	24	0.48	0.28	0.68
United States	26	36	28	5	63	15	0.48	0.40	0.77

Table 1: Summary of fiscal systems and inequalities in the US and and in France. Total taxes, public debt B and public spending G in percentage of GDP; tax rates τ^K, τ^L and τ^c in percent; Gini indices unitless.

The first column reports the total mandatory levies as a share of GDP for the two countries. Following the literature, we report the decomposition of these total levies into three components: capital tax, labor tax, and consumption tax. Since Mendoza et al. (1994), this decomposition is widely used to compare the tax structure across countries (OECD, Eurostat). These three taxes are reported in columns 2–4. The second column shows the implicit capital tax, calculated as tax receipts on capital income divided by the capital stock. The third column provides the same statistic for the labor tax and is computed as the tax receipts on labor income divided by the aggregate labor supply. The fourth column reports the implicit tax on consumption.

We can observe that overall taxes are 50% higher in France than in the US. Although capital taxes are very close in both countries, the labor and consumption taxes differ significantly. The difference in labor tax partly stems from the financing of the French welfare system, which covers public pensions, unemployment benefits, health care, and family allowances. It mostly relies on social contributions based on the wage bill, which are considered as labor tax. Regarding consumption tax, it is much higher in France compared to the US, although this high value is comparable to those in other European countries. Tax revenues are used to finance public spending, which includes both public consumption and public investment. Public spending, as a share of GDP, (column G) is approximately 60% higher in France than in the US. This difference is partly explained by larger investments in public infrastructure in France. Despite different levels of taxation and public spending, the public debt-to-GDP ratio appears to be comparable in France and in the US at around 60%.

We also report in Table 1 the evolution of income inequality before and after taxation. We proxy income inequality using the average Gini index between 1995 and 2007 (included), as reported in the OECD Income Inequality Database.[4] Note that the Gini indices barely vary over the period, and the picture would not have been different if we had reported the 2007 data only. The before-tax Gini indices for income are roughly similar in France and in the US. This value for France stems from the accounting of the (high) public pensions in France, which are counted as transfers and not as income. Consequently, this contributes to increasing the before-tax inequalities. However, the after-tax Gini indices are very different in the two countries, which is a consequence of the high transfers to households in France. While redistribution diminishes the Gini index for income by less than 10 points in the US, the reduction is twice as large in France and amounts to 20 points.

The last column reports the Gini index for wealth. The data for France come from the Household Finance and Consumption Survey (HFCS) for the 2010 wave, which is the closest

[4]See `https://stats.oecd.org/index.aspx?queryid=66670`.

wave to our benchmark years. We have checked that the Gini index remains highly similar in the other waves. The wealth Gini index for the US is taken from the PSID in 2006.[5] As is standard, wealth inequalities in each country are higher than for income. The wealth Gini index in each country is approximately 30 points higher for wealth than for post-tax income. The comparison for wealth between the US and France yields a result similar to that of the post-tax income did: The US value is approximately 10 points higher than the French one. It confirms that inequalities are more pronounced in the US than in France.

Although the results in Table 1 consider a linear tax for labor (column τ^L), the labor income tax scheme is actually progressive both in France and in the US. Comparing the progressivity of labor income tax across countries is challenging due to the complex tax schedules and deductions that are specific to each country. One approach to make this comparison tractable is to use a parametric form for the tax function. We follow the literature (e.g., Benabou, 2002 and Heathcote et al., 2017) and consider a log-linear functional form the labor tax:

$$\text{Tax:} \qquad T(I_c) = I_c - \kappa I_c^{1-\tau}, \tag{1}$$

$$\text{Disposable income:} \qquad D(I_c) = \kappa I_c^{1-\tau}, \tag{2}$$

where I_c is the labor income of the country c, τ the level of progressivity, and κ the average level of taxation. Notice that the higher the τ, the more progressive the tax system.

We use the Luxembourg Income Study (LIS) database for France and the US in 2005 to estimate the tax progressivity for labor income. We restrict our attention to the heads of households and their spouses aged between 25 and 60 who were employed. We define *labor income* as the sum of wage income, self-employment income, and private transfers. Using the estimates of the capital tax from Trabandt and Uhlig (2011), we can deduce from the capital income the amount of capital income tax. We then subtract from the total income tax (provided by the LIS) amount the capital tax amount, which allows us to obtain an estimated amount of the labor income tax. We finally define the *disposable income* as the *labor income* minus the *labor income tax* amount.

Using these data, we estimate the labor tax progressivity by regressing the log of disposable income on the log of labor income – which corresponds to the log of equation (2). Table 2 reports our estimation results for France and the US: $\hat{\tau}$ is the estimated labor tax progressivity and SE the associated standard error. France has a much more progressive labor tax than the US. Our estimate of progressivity for the US is 0.16, which closely aligns with values used in the literature. Our value is lower than the 0.181 value estimated by Heathcote et al. (2017), as we solely focus on estimating the progressivity of labor income and did not consider the progressivity of labor and capital income combined.

[5] In the 2007 SCF, the wealth Gini index was found to be 0.78, which is very close to the PSID value.

	$\hat{\tau}$	SE	Obs.	R^2
France	0.23	0.0056	5289	0.855
United States	0.16	0.0019	38111	0.942

Table 2: Estimates $\hat{\tau}$ of the progressivity of the labor income tax in the US and and in France for 2005. We regress the log of equation (2) using the LIS database: SE is the standard error of $\hat{\tau}$, Obs. is the number of observations, R^2 is the R^2 of this regression.

We will use the elements of Table 1 and the progressivity of labor income estimate in Table 2 to calibrate our heterogeneous-agent model below – and in particular the social weights.

3 A Bewley theory of the SWF: Some initial definitions

We discuss our aggregation theory of the SWF in a simple environment that allows us to abstract from complex algebra and heavy notation. We first explain how we construct individual welfare functions (IWFs), which are then aggregated to form the social welfare function (SWF) (Sections 3.1 and 3.2). We then present how the SWF and the IWFs can be identified from observed allocations (Section 3.3). We also introduce public finance concepts and also discuss how they relate to our identification strategy. Finally, we discuss the possible interpretations of these estimations in terms of political and social justice concepts, such as the Utilitarian, Egalitarian, and Libertarian principles (Section 3.4).

3.1 The setup

We consider a one-period one-good economy. The unique good is a final consumption good, over which agents have preferences. These preferences are represented by a utility function u, which is assumed to satisfy standard properties: $u' > 0$, $u'' < 0$, and $u'(0) = \infty$. We furthermore assume that $u > 0$. We need this assumption for the combination of weighted utility functions to be well-behaved. The economy is populated by two types of agents, whose population size is normalized to 1. Each type is in equal share 1/2. Agents types only differ along their endowment. We denote by $y_1 > y_2$ the two endowment values.

A benevolent planner has the objective to choose the best allocation subject to a feasibility constraint. Allocations will be ranked according to a SWF whose construction is detailed below. The feasibility constraint reflects the fact that the planner can transfer resources across type 1 and type 2 agents, but with a quadratic redistribution cost. This cost aims at capturing all distortions generated by distribution and is scaled by a parameter $\kappa > 0$. Implementing a consumption c_i for an agent receiving the endowment y_i involves a destruction of resources equal to $\frac{\kappa}{2}(c_i - y_i)^2$. Focusing on the symmetric equilibrium where all agents of the same type i receive the same consumption c_i, the feasibility constraint can be written as:

$$\sum_{i=1}^{2} \left(c_i + \frac{\kappa}{2}(c_i - y_i)^2 \right) \leq \sum_{i=1}^{2} y_i. \tag{3}$$

To consider meaningful solutions, we assume that the redistribution cost is not too high and verifies $\kappa y_i < 1$, which formally guarantees interior solutions.

3.2 Individual and social welfare functions

Our construction of the SWF from individual ethical preferences (IWFs) proceeds in three steps – that we will replicate in the general case of Section 4: (i) the subjective valuation of each agent for the welfare of others; (ii) the representation of ethical preferences which we will call Individual Welfare Functions (IWFs); (iii) the Social Welfare Function (SWF) representing the planner's preferences.

The Individual Welfare Function (IWF). In our static setup, $u(c_i)$ is the utility of agent i for the consumption c_i. However, agent i also has their own view of how their welfare and this of others should be accounted for by the planner. We model this subjective valuation of the utility of others agents by a loading factor that weights the individual utility. Denoting by $\tilde{\omega}_{ij}$ the weight of agent i for agent j, the subjective valuation by agent i of the welfare of agent j is $\tilde{V}_{ij} = \tilde{\omega}_{ij} u(c_j)$.[6]

We then assume that the ethical preferences of agent i are built as the aggregation of their perception of the welfare of other agents. The IWF of agent i representing their ethical preferences is thus the weighted sum of subjective valuations by agent i over the two agents' types: $IWF_i := \frac{1}{2}\tilde{V}_{i1} + \frac{1}{2}\tilde{V}_{i2}$, or:

$$IWF_i = \frac{1}{2}\tilde{\omega}_{i1} u(c_1) + \frac{1}{2}\tilde{\omega}_{i2} u(c_2). \tag{4}$$

The economy thus features heterogeneity in ethical preferences: There are two types of ethical preferences, as there are two income levels. Types thus capture the heterogeneity both in the endowments and in the IWFs. This is in line with empirical studies, that have indeed shown heterogeneity in ethical preferences and that this heterogeneity was partly driven by social position. See Gaertner and Schokkaert (2012) and Stantcheva (2021) for a recent analysis based on large US surveys.

This general case encompasses *self-interested* agents, who only care about their own welfare.[7] The IWF of self-interested agents is proportional to their individual utility: $IWF_i = \lambda_i u(c_i)$ for some $\lambda_i > 0$. This corresponds to the weights $\tilde{\omega}_{ij} = \lambda_i \times 1_{i=j}$ – with $1_{i=j} = 1$ if $i = j$ and 0 otherwise.

Two remarks are worth mentioning regarding the generality of the weights $(\tilde{\omega}_{ij})_{i,j=1,2}$

1. *Intensity of preferences.* We do not impose any weight normalization, neither that they sum to 1 ($\sum_j \tilde{\omega}_{ij} = 1$) nor that $\tilde{\omega}_{ii} = 1$. Since we aggregate IWFs together, the IWF weights have indeed both an ordinal and a cardinal meaning. Their normalization is not innocuous, as it would remove the possible heterogeneity in the intensity of ethical preferences (using the wording of Arrow, 1951).

2. *Possibility of spitefulness or discrimination.* We do not restrict the sign of the weights $(\tilde{\omega}_{ij})_{i,j=1,2}$, which are allowed to be negative. In this case, the IWF of an individual may be

[6] Note that to avoid imposing an implicit normalization, we do not impose $\tilde{\omega}_{ii} = 1$. See also below our discussion about the intensity of preferences.

[7] We avoid the word "selfish" or "rational" agents, as they are too negatively or positively connoted. Sen (1977) uses the word self-seeking for the same idea.

negatively affected by the utility of other agents: the consumption of some other agents can be perceived as a negative externality by certain agents. Such externalities on individual preferences (and not only on ethical preferences as here) have been modeled in asset pricing or macroeconomics to reflect the idiom that households want "to keep up with the Joneses" (see Abel, 1990 or Campbell and Cochrane, 1999 among others). Such externalities have also received support in experimental studies (see Fehr et al., 2013 among others), where such a behavioral trait has been called spitefulness.[8] Negative weights in welfare functions could also reflect a possible discrimination of some agents' types (Piacquadio, 2017).

The Social Welfare Function (SWF). Finally, we construct the SWF as the weighted aggregation of the IWFs. Indeed, following the political economy literature, we assume that agents may differ in their political ability to influence the planner in their policy implementation. This heterogeneity in the agents' influence may result from institutional design, lobbying activities, or voting rules. We capture it by political weights $(\omega_{P,i})_i$ which loads IWFs as a function of agent's type. We will use the political economy literature to estimate these weights. We then define the SWF as the aggregation of IWFs weighted by political weights:

$$SWF = \omega_{P,1} IWF_1 + \omega_{P,2} IWF_2. \tag{5}$$

Using the expression (4) of IWFs, we obtain the following expression for the SWF:

$$SWF = \omega_1 u(c_1) + \omega_2 u(c_2), \tag{6}$$

where we have defined the SWF weights as follows, for some $\omega > 0$:

$$\omega_1 := \frac{\omega}{2}(\omega_{P,1}\tilde{\omega}_{11} + \omega_{P,2}\tilde{\omega}_{21}), \tag{7}$$

$$\omega_2 := \frac{\omega}{2}(\omega_{P,1}\tilde{\omega}_{12} + \omega_{P,2}\tilde{\omega}_{22}). \tag{8}$$

While the IWFs weights have a cardinal interpretation, this is not the case of the SWF weights. We can thus without loss of generality choose the constant ω so as to normalize the sum of weights to 1 ($\omega_1 + \omega_2 = 1$).

Our construction of SWF weights embeds standard cases, such as the Utilitarian SWF. It corresponds to self-interested agents: $\tilde{\omega}_{ij} = 1_{i=j}$, with identical preference intensity. With constant political loading factors: $\omega_{P,i} = 1/2$, we obtain that the SWF weights are identical and equal to: $\omega_1 = \omega_2 = \frac{1}{2}$ – where we have set $\omega = 2$ in (7)–(8) for normalization purposes. The resulting SWF is Utilitarian.[9] Moreover, SWF weights are not restricted to be positive and the aggregation procedure could be end up with negative SWF weights. In this simple setup, this implies that the construction does not ensure that the planner only chooses Pareto-optimal allocations. See also Section 3.4 for a lengthier discussion of these aspects.

[8] Fehr and Schmidt (2006) describe this behavior as follows: "A spiteful person always values the material payoff of relevant reference agents negatively. Such a person is, therefore, always willing to decrease the material payoff of a reference agent at a personal cost to himself".

[9] Even with heterogeneous preference intensity, $\tilde{\omega}_{ij} = \lambda_i 1_{i=j}$, with $\lambda_1 \neq \lambda_2$, we can recover the Utilitarian SWF. Political weights must offset the preference intensity: $\omega_{P,i} = \frac{\lambda_i^{-1}}{\lambda_1^{-1}+\lambda_2^{-1}}$. We then obtain with $\omega = \lambda_1^{-1} + \lambda_2^{-1}$ that SWF weights are identical: $\omega_1 = \omega_2 = \frac{1}{2}$.

3.3 Inverse optimal approach: Identifying the Welfare Functions

The identification of SWF weights. Generally speaking, the inverse optimal approach consists in identifying social preferences from the observed allocation and available fiscal tools, assuming that the latter are optimally set by the planner. In our setup, the fiscal system allows the planner to directly choose the allocation (c_1, c_2) that maximizes the aggregate welfare, represented by the SWF of equation (6), subject to the resource constraint of equation (3). For a given pair of SWF weights (ω_1, ω_2), the optimal allocation is characterized by the following FOC:

$$\frac{\omega_1 u'(c_1)}{1+\kappa(c_1-y_1)} = \frac{\omega_2 u'(c_2)}{1+\kappa(c_2-y_2)}, \tag{9}$$

together with the constraint (3). Note that the condition $\kappa y_i < 1$ ensures that the FOC is well defined for all consumption levels.

The intuition for this relationship can be provided using concepts of public finance. To clarify this link, we rewrite the planner's program assuming that the planner chooses individual lump-sum taxes $(t_i)_i$. The Lagrangian of the planner can be written as $\mathcal{W} + \mu\mathcal{B}$, where $\mathcal{W} := \omega_1 u(y_1 - t_1) + \omega_2 u(y_2 - t_2)$ is the SWF (6) expressed with taxes, $\mathcal{B} := \sum_{i=1}^{2}(t_i - \frac{\kappa}{2}t_i^2)$ is the resource constraint (3) as a function of taxes, and μ is the associated Lagrange multiplier. The FOC associated to the choice of t_i can be written as follows:

$$\mu = \frac{\partial \mathcal{W}}{\partial c_i} \frac{-\frac{\partial c_i}{\partial t_i}}{\frac{\partial \mathcal{B}}{\partial t_i}}. \tag{10}$$

This equation states that the planner equalizes the marginal benefit for the planner's finances of raising t_i to the marginal cost for agent i of financing this marginal resource. The marginal benefit is simply the Lagrange multiplier of the planner's resource constraint. The marginal cost involves two terms.

The first one, $-\frac{\partial c_i}{\partial t_i} \Big/ \frac{\partial \mathcal{B}}{\partial t_i}$, measures how much the consumption of agent i is affected by the financing of the marginal planner's resource by the tax t_i. The quantity $\frac{\partial \mathcal{B}}{\partial t_i} = 1 - \kappa t_i$ includes includes the tax base (equal to 1 here) and the financial externality related to the destruction of resources, $-\kappa t_i$. Since $-\frac{\partial c_i}{\partial t_i} = 1$, financing the marginal planner's resource will decrease the consumption of agents i by $-\frac{\partial c_i}{\partial t_i} \Big/ \frac{\partial \mathcal{B}}{\partial t_i} = 1/(1+\kappa(c_i - y_i))$. We denote this term $MVPF_i$ and call it the *marginal value of public fund* following Finkelstein and Hendren (2020) and Hendren and Sprung-Keyser (2020).

The second term measures $\frac{\partial \mathcal{W}}{\partial c_i}$ how much social welfare is affected by a variation of the consumption of agents i. In the absence of welfare externality of consumption, only the welfare of agent i is affected and we have $\frac{\partial \mathcal{W}}{\partial c_i} = \omega_i u'(c_i)$. Following the literature (e.g., Ferey et al., 2024 among many others), we call this term the *social marginal welfare weight* attributed by the planner to agents i. We denote it as $SMWW_i$.

Overall, the extra resource of the planner financed by agent i through t_i implies a consumption cut of $MVPF_i$ units for agents i, and a welfare impact equal to $SMWW_i \times MVPF_i$. This latter quantity can thus be interpreted as the *bang for the buck* of one unit of resources spent by the planner for agents i – following again the denomination of Finkelstein and Hendren (2020) and Hendren and Sprung-Keyser (2020).

The interpretation of equation (9) is then rather simple. The planner sets transfers between the two agents' types up to the point where the planner equalizes the bangs for the buck of the two agents: the planner is indifferent between obtaining one extra unit of resources from agents 1 or from agents 2. Should this not hold, the allocation could not be optimal: Agents with the higher bang for buck should receive resources, at the expense of those with the lower bang for the buck. Furthermore, from (10), the bangs for the buck of the two agents are also equal to the marginal benefit of relaxing the resource constraint (3). Equation (9) can thus be rewritten as:

$$SMWW_1 \times MVPF_1 = SMWW_2 \times MVPF_2 = \mu. \tag{11}$$

Saez and Stantcheva (2016) have generalized this marginal approach by allowing to consider social marginal weights that do not derive from an explicit SWF. The weights can adopt general expressions, such as non linear effects or dependencies in endogenous variables other than consumption. These weights are called *generalized social marginal welfare weight*.[10]

The inverse optimal approach still relies on the FOC (9), but takes a different perspective. Instead of deducing the allocation (c_1, c_2) from the weights (ω_1, ω_2), the weights are computed from the allocation. Formally, for a given allocation (c_1, c_2) satisfying the resource constraint (3), the FOC (9) can be written as:

$$\frac{\omega_1}{\omega_2} = \frac{u'(c_2)}{u'(c_1)} \frac{1 + \kappa(c_1 - y_1)}{1 + \kappa(c_2 - y_2)}, \tag{12}$$

which determines the pair of weights (ω_1, ω_2) with the normalization constraint. Rather than the SWF weights ω_i, we could also compute from (9) the social marginal welfare weights, $\omega_i u'(c_i)$. As we discuss in Section 3.4, focusing on SWF weights is better suited for our analysis, as it allows us to direct qualify SWF and offer possible interpretations in terms of political and ethical terms.

The identification of IWF weights. Most of the analysis in the quantitative model of Section 6 involves the estimation of the SWF. However, it also insightful to derive the IWFs that are consistent with the estimated SWF, because it allows us to understand the underlying heterogeneity in social perceptions. To do so, we implement the following strategy. First, we consider measures of the political weights $(\omega_{P,i})_i$ of each group, using insights from the political economy literature. Second, in the absence of individual-level information, we identify the weights of IWFs as those, which are the closest to the self-interested ones, while being consistent with the estimated SWF weights. See Section 4.4 for a formal presentation.

Other benchmark IWFs could easily be considered, but the self-interested ones appear as standard in the economic or political economy literature (see Acemoglu, 2010 for various models of this type). The gain of this strategy is to make explicit the identifying assumptions and to allow us to derive closed-form expressions for the IWF weights.

[10] The GSMWW approach allows for including a wide variety of political or moral motives. However, when the Pareto principle is also imposed, Sher (2024) has shown that the GSMWW approach may generate inconsistencies in the ranking of fiscal schedules.

3.4 Interpretations of Welfare Functions

Welfare weights are known offer a possible interpretation in terms of social choice and moral philosophy, when the allocation is known (Saez and Stantcheva, 2016). We provide below the definitions that will be used in the quantitative exercise of Section 6.

- *Utilitarian.* Agents of type i will be said to be Utilitarian if they equally weight all agents: $\tilde{\omega}_{i1} = \tilde{\omega}_{i2}$. The Utilitarian planner ($\omega_1 = \omega_2$) will implement the consumption levels $c_1 > c_2$ if $\kappa > 0$ (recall that $y_1 > y_1$). The Utilitarian planner thus accepts some inequality among agents due to the distributional cost.

- *Egalitarian.* Egalitarian agents think that economic inequality isn't justified. Formally, this implies putting a greater weight on poorer agents: $\tilde{\omega}_{i1} < \tilde{\omega}_{i2}$. Compared to a Utilitarian planner, the Egalitarian planner will reduce the consumption of the richest agents (type 1) and increase the consumption of poorest agents (type 2). The planner thus reduces inequality at the cost of a lower total consumption.

- *Libertarian.* Libertarian agents think that agents deserve what they have. This formally corresponds to a higher weight on richer agents: $\tilde{\omega}_{i1} > \tilde{\omega}_{i2}$. The Libertarian planner will choose an allocation implementing a higher inequality but a higher total consumption than the Utilitarian planner. The resulting allocation is thus closer to the initial income distribution and generates lower distributional costs.

In a two-type economy, these three cases correspond to a partition of the set of welfare weights. Each agent type can only in one of the three above situations. This is also the case for the planner. Note that even if the agent types belong to different categories, the planner will belong to exactly one of those categories – depending on the combination of political weights and preference intensity. However, such a partitioning does not hold with more than two types of agents. We discuss such cases in the general model of Section 4.

What about Pareto deviations? Our SWF construction is not restricted to Pareto-optimal SWF. In our simple setup, non-Pareto optimal allocations correspond to negative SWF weights. In more general settings, this can happen even with positive SWF weights. This is for instance the case in a framework featuring individual risk and a Libertarian planner. In that case, an insurance mechanism could be ex-ante individually optimal for all agents, while not being chosen by the planner. Indeed, a Libertarian planner would be reluctant to favor redistribution. Thus, the SWF allocation may not be Pareto optimal, as it could prevent insurance mechanisms. This is not specific to Libertarian planner. An Egalitarian planner could choose an allocation that reduces so much the inequality that it would come at the expense of some agents. Indeed, since Sen (1970), it is known that the SWF will not necessarily fulfill the Pareto principle if the planner also care for other factors (ethical, moral, or political) than the individuals' utility. More precisely, Kaplow and Shavell (2001) have shown that the Pareto principle does not hold when the planner departs from welfarism, i.e., when the SWF does not only include the agents' utility. There is therefore an incompatibility between Pareto principle and the generality of the SWF. One of our paper objective is to propose a rationalization of observed fiscal systems and discuss political implications. As a consequence, we will relax the assumption of welfarism and will not require the Pareto principle to be fulfilled by the estimated SWFs. We alternatively consider a weaker restriction, which is that the aggregate welfare cannot decrease when increasing the welfare of any agents. Loosely speaking, this is akin to "everybody counts". This means assuming positive SWF weights: $\omega_1, \omega_2 \geqslant 0$ (see Definition 1 in the intertemporal case).[11]

4 A Bewley theory of the SWF: The intertemporal case

We extend the previous construction of the SWF to a general intertemporal framework, which is our first contribution. Our approach relies on the sequential representation of the heterogeneous agent model, which is the most suitable for our normative analysis.[12] We present the construction of the SWF anticipating the model used in the quantitative section below.

4.1 The setup

We now consider an infinite-horizon model with incomplete financial markets. Time is discrete, indexed by $t \geq 0$. The economy is populated by a continuum of size 1 of ex-ante identical agents.

Risk structure. Idiosyncratic risk is modeled as an uninsurable idiosyncratic labor productivity shock y_t that can take Y distinct values in the finite set $\mathcal{Y}$. The productivity risk follows a first-order Markov chain with transition matrix $(\Pi_{yy'})_{y,y' \in \mathcal{Y}}$. This matrix is assumed to be irreducible and aperiodic, which ensures that it admits a unique stationary distribution denoted as $(\pi_y)_{y \in \mathcal{Y}}$, normalized such that $\sum_{y \in \mathcal{Y}} \pi_y = 1$. We denote by $y^t = \{\ldots, y^t_{t-1}, y^t_t\}$ a one-sided infinite sequence of elements of $\mathcal{Y}$, corresponding to an history of productivity levels up to date t. We denote the set of such histories by $\mathcal{Y}^\infty$. Since we will need to consider the evolution of histories from one period to another, we keep time subscripts for histories. To keep notation simple, we will use for an history $y^t \in \mathcal{Y}^\infty$, the following notation: (i) $y^t_\tau \in \mathcal{Y}$ is the productivity level at date $\tau \leq t$ in history y^t; (ii) $y^{s,t}$ is the truncation of y^t at date $s \leq t$ – such that y^t and $y^{s,t}$ coincide up to s: $y^{s,t}_\tau = y^t_\tau$ for all $\tau \leq s$. We will use decorator to clearly distinguish possible different histories: $\tilde{y}^t$ and y^t can be different at any date.

Initial distribution. To simplify the notation below, we make two assumptions about the initial distribution: (i) all agents start the economy with an initial infinitely-long history belonging to $\mathcal{Y}^\infty$; (ii) agents are initially endowed with a wealth that is only function of their history. The assumption about wealth encompasses, among others, the case where all agents have the same wealth or the steady-state wealth. Since we focus on steady-state distributions, of wealth in particular, this simplification is at no cost in our environment.

[11]In our simple setup, this weaker restriction is equivalent to the Pareto principle, but this is not the case in more general setups.

[12]Some analyses (e.g., Chang et al., 2018) have considered social weights depending on endogenous variables, such as consumption or wealth. This is a possible source of inconsistencies and multiple equilibria. Indeed, the planner should in that case consider how the social weights change with the allocation.

The sequential representation. There is a mathematical subtlety in infinite-horizon models, as the set of histories has the cardinality of the continuum (it is neither finite nor countable). This explains why the probability space over the set of histories involves a general measure. We thus construct a probability space over the set of all histories denoted by $(\mathcal{Y}^\infty, \mathcal{F}, \mu)$, where $\mathcal{F}$ is a relevant σ-algebra and μ is a measure (see Appendix A.1). In words, for any set of histories $B \in \mathcal{F}$, $\mu(B) \geq 0$ is the measure of agents currently experiencing an history $y^t \in B$.[13] As the population size of agents is normalized to 1, we furthermore have $\int_{y^t \in \mathcal{Y}^\infty} \mu(dy^t) = 1$.[14]

We also need to define transitions across histories. Consider two histories $y^{t+1}, \tilde{y}^t \in \mathcal{Y}^\infty$. The probability to switch from history $\tilde{y}^t$ in the current period to another history y^{t+1} in the next period is simply the probability to switch from state $\tilde{y}^t_t$ to state y^{t+1}_{t+1} if $y^{t,t+1}$ and $\tilde{y}^t$ are equal and 0 otherwise. We denote this conditional probability $\mu_1(\cdot|\tilde{y}^t)$, formally defined as: $\mu_1(dy^{t+1}|\tilde{y}^t) = \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(dy^{t,t+1})$, where $\delta_{\tilde{y}^t}$ is the Dirac delta function in $\tilde{y}^t$.[15]

We then define by induction the probability to switch from history $\tilde{y}^t$ to another history y^{t+s} s periods ahead as: $\mu_s(dy^{t+s}|\tilde{y}^t) = \Pi_{y^{t+s}_{t+s-1} y^{t+s}_{t+s}} \mu_{t-1}(dy^{t+s-1,t+s}|\tilde{y}^t)$, or: $\mu_s(dy^{t+s}|\tilde{y}^t) = \prod_{k=0}^{s-1} \Pi_{y^{t+s}_{t+k} y^{t+s}_{t+k+1}} \times \delta_{\tilde{y}^t}(dy^{t,t+s})$. In words, switching from $\tilde{y}^t$ to y^{t+s} imposes that the two histories coincide up to period t and then involves the cumulative probability to successively experience the states from y^{t+s}_{t+1} to y^{t+s}_{t+s}.

Individual intertemporal welfare. For a given allocation, we denote by $U(y^t)$ the period utility of an agent having history y^t. To lighten notation, we choose not to make the dependence in the allocation explicit. For instance in the case of a utility depending on private consumption, and labor supply (as in our quantitative application), $U(y^t) := u(c(y^t)) - v(l(y^t))$, where: $c : \mathcal{Y}^\infty \to \mathbb{R}^+$ and $l : \mathcal{Y}^\infty \to \mathbb{R}^+$ are policy functions determining consumption and labor as a function of individual history. We still assume that U is always positive, which ensures, as in the simple case, that the combination of weighted utility functions is well-behaved.

The intertemporal welfare in period t of an agent with history y^t is assumed to be separable in time and of the expected-utility type. It is thus defined as the discounted sum over all future dates of expected period utilities. Formally, the intertemporal utility $V(y^t)$ is:

$$V(y^t) = \sum_{s=0}^{\infty} \beta^s \int_{\tilde{y}^{t+s} \in \mathcal{Y}^\infty} U(\tilde{y}^{t+s}) \mu_s(d\tilde{y}^{t+s}|y^t). \tag{13}$$

As we consider steady-state allocation, the intertemporal welfare can be written recursively as:

$$V(y^t) = U(y^t) + \beta \mathbb{E}_{\tilde{y}^{t+1}} \left[V(\tilde{y}^{t+1}) | y^t \right], \tag{14}$$

where $\mathbb{E}_{\tilde{y}^{t+1}} \left[V(\tilde{y}^{t+1}) | y^t \right] = \int_{\tilde{y}^{t+1} \in \mathcal{Y}^\infty} V(\tilde{y}^{t+1}) \mu_1(d\tilde{y}^{t+1}|y^t)$ is a conditional expectation. See Appendix A.3 for a proof.

[13] An history is an event of measure zero in $\mathcal{F}$. Therefore, every equality that holds for all history y^t should be understood as holding almost surely.

[14] In finite time, we would write $\sum_{y^t \in \mathcal{Y}^t} \mu_t(y^t)$ instead of $\int_{y^t \in \mathcal{Y}^\infty} \mu(dy^t)$. All intuitions of the finite-time representation are valid.

[15] The function μ_1 is a measure and verifies the standard properties of a conditional probability: $\mu_1 \geq 0$, $\int_{y^{t+1} \in \mathcal{Y}^\infty} \mu_1(dy^{t+1}|\tilde{y}^t) = 1$ and $\int_{\tilde{y}^t \in \mathcal{Y}^\infty} \mu_1(dy^{t+1}|\tilde{y}^t)\mu(d\tilde{y}^t) = \mu(dy^{t+1})$. See the proofs in Appendix A.2.

4.2 Constructing the SWF

We extend the SWF construction of Section 3 to the current intertemporal framework. As previously, our construction involves three steps. The first step is slightly different than in the simple case. In the simple framework, agents of type i had their own subjective valuation of how the utility of agents of type j should be valued by the planner. Since histories are the sole source of heterogeneity among agents, the perception of others' situation thus relies on the perception of other histories. This results in a subjective valuation by agents y^t of the utility of other agents with history $\tilde{y}^t$. The second and third steps are more similar to their counterpart of the simple model. The second step is to aggregate the perception of other histories' over the entire distribution μ of histories, which yields the Individual Welfare Function (IWF) of agents with history y^t. The third and final step is to construct the Social Welfare Function (SWF) as the weighted sum of individual IWFs over all histories y^t. We now present this aggregation more formally.

Step 1: Constructing the subjective valuation of the utility of another agent. We consider two groups of agents characterized by their histories $y^t \in \mathcal{Y}^\infty$ and $\tilde{y}^t \in \mathcal{Y}^\infty$ at some date t; the allocation is still considered as given. Given our previous assumption about individual preferences, the tilde agents value in $s \geq 0$ periods the allocation of any history $\hat{y}^{t+s}$ as $U(\hat{y}^{t+s})$ (whether $\hat{y}^{t+s}$ is a possible successor of $\tilde{y}^t$ or not). However, the non-tilde agents may possibly have a different perception of how the allocation history $\hat{y}^{t+s}$ should be valued. Non-tilde agents assign to the utility $U(\hat{y}^{t+s})$ of the tilde agents a corrective factor, denoted by $\hat{\omega}(y^t, \hat{y}^{t+s})$, such that $\hat{\omega}(y^t, \hat{y}^{t+s})U(\hat{y}^{t+s})$ corresponds to the valuation of history $\hat{y}^{t+s}$ by the non-tilde agents. Summing over all future periods and all future histories the discounted valuations $\hat{\omega}(y^t, \hat{y}^{t+s})U(\hat{y}^{t+s})$ – with the proper conditional probabilities – yields the valuation of the utility of the tilde agents by the non-tilde agents. Denoting by $\hat{V}(y^t, \tilde{y}^t)$ this subjective valuation, we obtain:

$$\hat{V}(y^t, \tilde{y}^t) = \sum_{s=0}^{\infty} \int_{\hat{y}^{t+s} \in \mathcal{Y}^\infty} \beta^s \hat{\omega}(y^t, \hat{y}^{t+s}) U(\hat{y}^{t+s}) \mu_s(d\hat{y}^{t+s}|\tilde{y}^t), \tag{15}$$

which is a direct modification of the utility $V(\tilde{y}^t)$ with the inclusion of the weights $\hat{\omega}(y^t, \hat{y}^{t+s})$.

Step 2: Constructing the Individual Welfare Function (IWF). The IWF of the agents with history y^t is then constructed as the aggregation of their subjective valuation $\hat{V}(y^t, \tilde{y}^t)$ over possible histories $\tilde{y}^t$. Formally:

$$IWF(y^t) = \int_{\tilde{y}^t \in \mathcal{Y}^\infty} \hat{V}(y^t, \tilde{y}^t) \mu(d\tilde{y}^t). \tag{16}$$

The IWF is a representation of the ethical preferences of agents with history y^t. It represents how the agents y^t think the welfare of all other agents should be accounted for by the planner. As in the simple model, we do impose any weight normalization at this stage.

Step 3: Aggregating IWFs to obtain the Social Welfare Function. We assume that the planner observes the IWFs in the population and aggregates them all depending on the weights assigned to each agent. Not all agents have the same importance for the planner and they differ along what we call their political weights – as in the simple model. Formally, the IWF of agents with history y^t will be assigned by the planner the weight $\omega_P(y^t)$. This weight is a shortcut for the importance of agents with history y^t in the political process and hence in their ability to have their own IWF accounted for by the planner. Formally:

$$SWF = \int_{y^t \in \mathcal{Y}^\infty} \omega_P(y^t) IWF(y^t)\mu(dy^t). \tag{17}$$

Special cases. To illustrate our construction, we now consider special cases.

The first case is when agents identically value other histories. Formally, the weights of agent y^t are the same for all histories $\hat{y}^{t+s}$: $\hat{\omega}(y^t) := \hat{\omega}(y^t, \hat{y}^{t+s})$, with a slight abuse of notation. In that case, the ethical preferences can be shown to be represented by an IWF that is proportional to the utilitarian SWF: $IWF(y^t) = \hat{\omega}(y^t) \int_{\tilde{y}^t \in \mathcal{Y}^\infty} V(\tilde{y}^t)\mu(d\tilde{y}^t)$. All agents have the same weights in the agents' ethical preferences. Consequently, there is no disagreement in the population for the ordering of allocation. The SWF reflects this absence of disagreement and is also proportional to the utilitarian SWF: $SWF = \left(\int_{\tilde{y}^t \in \mathcal{Y}^\infty} \omega_P(\tilde{y}^t)\hat{\omega}(\tilde{y}^t)\mu(d\tilde{y}^t)\right) \int_{y^t \in \mathcal{Y}^\infty} V(y^t)\mu(dy^t)$.[16]

Second, we consider the so-called self-interested agents, who only care about the histories they can possibly experience. Formally, the weights of an agent with history y^t will be zero for histories that are not possible continuations of y^t: $\hat{\omega}(y^t, \hat{y}^{t+s}) := \delta_{y^t}(\hat{y}^{t,t+s})\hat{\omega}(y^t)$, with a slight abuse of notation again. In that case, the ethical preferences of agents with history y^t are identical to their individual preferences and their IWF is proportional to their intertemporal utility: $IWF(y^t) = \hat{\omega}(y^t)V(y^t)$. This illustrates that these agents only care about themselves, which justifies our denomination of "self-interested". In that case, the SWF is equal to a weighted sum of individual intertemporal utilities: $SWF = \int_{y^t \in \mathcal{Y}^\infty} \omega_P(y^t)\hat{\omega}(y^t)V(y^t)\mu(dy^t)$, which is a weighted additive SWF. It reduces to a utilitarian SWF when the weight product $\omega_P(\cdot)\hat{\omega}(\cdot)$ is constant.

4.3 Properties of the SWF

An explicit expression of the SWF. We state the following proposition.

Proposition 1 *The SWF (17) admits the following expression:*

$$SWF = \sum_{t=0}^{\infty} \int_{y^t \in \mathcal{Y}^\infty} \beta^t \omega(y^t) U(y^t)\mu(dy^t). \tag{18}$$

where the weights ω are given by:

$$\omega(y^{t+s}) = \int_{\tilde{y}^t \in \mathcal{Y}^\infty} \omega_P(\tilde{y}^t)\hat{\omega}(\tilde{y}^t, y^{t+s})\mu(d\tilde{y}^t). \tag{19}$$

The proof can be found in Appendix A.4. Proposition 1 provides a simple expression for the SWF. It states that we can find period weights ω depending on the current history such that the SWF expresses as the discounted sum over all dates and histories of the utility of that

[16]This property is known at least since Aiyagari (1995), to justify the use of the Utilitarian SWF under the veil of ignorance.

date and history, weighted by the factor ω. In other words, this twists the standard utilitarian SWF by weighting period utilities by a factor depending on the period history – the utilitarian SWF corresponding to a constant ω. The SWF weight $\omega(y^{t+s})$ in (19) can be interpreted as the "average" weight given to history y^{t+s} by all agents in the economy, where agents are weighted by their political leverage ω_P.

The sequential representation (18) of the SWF can also be written as a recursive representation: $SWF = \int_{y^t \in \mathcal{Y}^\infty} \omega(y^t)U(y^t)\mu(dy^t) + \beta \cdot SWF$, which can be seen as the extension of the recursive representation of the utilitarian SWF to history-dependent weights. This recursive representation is very simple because of our stationarity assumption. When considering the whole dynamics of the economy, the utility U_t is time-dependent (because of time-dependent allocation). The SWF representation is then: $SWF_t = \int_{y^t \in \mathcal{Y}^\infty} \omega(y^t)U_t(y^t)\mu(dy^t) + \beta \cdot SWF_{t+1}$. We use the latter representation when solving the Ramsey program.

Weight restriction. As discussed in the simple framework, we do not restrict the SWF to satisfy the Pareto principle. We do, however, impose a weaker restriction. To formally express this restriction, we need to make the dependence in the allocation explicit. We now denote the period utility $U : \mathcal{Y}^\infty \times \mathcal{A} \to \mathbb{R}$ and the SWF: $SWF : \mathcal{A} \to \mathbb{R}$, where $\mathcal{A}$ is the set of allocations. The period utility for an history y^t and an allocation A will be denoted by $U(y^t, A)$. For instance, in the case of our quantitative application, we denote $U(y^t, A) := u(c(y^t)) - v(l(y^t))$, where A is the pair of policy functions (c, l). We can state our result using this notation.

Definition 1 *A SWF $SWF : \mathcal{A} \to \mathbb{R}$, associated with a period utility $U : \mathcal{Y}^\infty \times \mathcal{A} \to \mathbb{R}$, is said to be element-wise monotone if for any two allocations A and A' such that $U(y^t, A) \geq U(y^t, A')$ for all y^t, we have $SWF(A) \geq SWF(A')$.*

This definition states that with a monotone SWF, if an allocation is in every period better (in the sense of the period utility) than another one, the former will always be preferred, in the sense of the SWF, to the latter. This property is similar to element-wise monotonicity for utility functions. Obviously, this is weaker than the Pareto principle, which would require A to be preferred to A' in the sense of the intertemporal utility, and not only of period utility.[17]

Proposition 2 *A SWF fulfills element-wise monotonicity iff the weights ω defined in (19) are nonnegative.*

The proof can be found in Appendix A.5. Our quantitative estimations may impose the positivity of weights, which corresponds to an element-wise monotone SWF. This means that the SWF cannot increase if the welfare of one agent is reduced: everybody (positively) counts.

4.4 Identification of the weights

The SWF expression in Proposition 1 is very general and does not easily lend itself to estimation. We thus introduce a tractability assumption that allows us to compute the weights in the SWF

[17] This explains why the planner may choose not to implement an insurance mechanism, even if this mechanism is individually desirable. Indeed, the mechanism implies higher utility in some states (typically the "bad" ones) and lower utility in others (typically the "good" ones). This does not verify our element-wise monotonicity of Definition 1.

and the IWFs. We assume that agents will the same current productivity level all value identically future histories and that this valuation actually only depends on the current productivity level of the history under consideration.[18] We make a similar assumption for political weights that are also supposed to only depend on the current productivity level. More formally:

Assumption A *There exist two functions, denoted with a slight abuse of notation* $\tilde{\omega} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ *and a function* $\omega_P : \mathcal{Y} \to \mathbb{R}$ *such that for all histories* $y^t, \tilde{y}^{t+s} \in \mathcal{Y}^\infty$*, we have:*

$$\begin{cases} \tilde{\omega}(y^t, \tilde{y}^{t+s}) := \hat{\omega}_{y_t^t, \tilde{y}_{t+s}^{t+s}}, \\ \omega_P(y^t) := \omega_{P, y_t^t}, \end{cases}$$

where we recall that y_t^t *is the current productivity level in history* y^t*.*

Assumption A reduces the dimensionality of loading factors, which are now defined on finite sets. Weights can now be interpreted as loading factors on productivity levels. To avoid heavy notation, we keep the same notation but use subscripts to denote the dependence in productivity level. The weights ω of equation (19) can also now be shown to verify for all $y \in \mathcal{Y}$: $\omega_y = \sum_{\tilde{y} \in \mathcal{Y}} \pi_{\tilde{y}} \omega_{P,\tilde{y}} \tilde{\omega}_{\tilde{y},y}$, where $\pi_{\tilde{y}}$ is recalled to be the share of agents with productivity $\tilde{y}$.

The identification of the weights proceeds in two steps: (i) the SWF weights $(\omega_y)_{y \in \mathcal{Y}}$ from the data; and (ii) the IWF weights $(\omega_{\tilde{y}y})_{y,\tilde{y} \in \mathcal{Y}}$ from the SWF weights and the data. Without loss of generality, we now impose for the identification a normalization constraint of the weights that are assumed to sum to 1: $\sum_y \pi_y \omega_y = 1$.

The SWF weights. Our identification strategy for the SWF weights consists in finding the weights for which a given fiscal system (typically the observed one) can be seen as the outcome of a Ramsey program. Formally, the FOCs of the Ramsey planner imply some linear constraints for the SWF weights $(\omega_y)_{y \in \mathcal{Y}}$. The number of constraints, n, depends on the number of instruments of the planner. In the general case, the system of constraints (the n linear ones and the normalization) is weakly underdetermined, which means that the number of productivity state $|\mathcal{Y}|$ is greater than the number of constraints, $|\mathcal{Y}| \geq n + 1$ (for any set X, $|X|$ is the cardinality of X).[19] There are thus $p = |\mathcal{Y}| - n - 1$ degree of freedom.

Our solution to handle the underdetermination is to follow Heathcote and Tsujiyama (2021) and to assume that the weights are a parametric function of y, with exactly $n + 1$ parameters. With mild assumptions on the functional forms, the weights are exactly and uniquely identified, as the solution of the non-linear system of constraints. We summarize it in the following definition.

Definition 2 *We consider as given a set of* $1 \leq n \leq |\mathcal{Y}| - 1$ *linear constraints represented by the matrix* $(L_{k,y})_{k=1,\ldots,n,y \in \mathcal{Y}}$ *and a set of parametric functions* $f_y : \mathbb{R}^{n+1} \to \mathbb{R}$ *(*$y \in \mathcal{Y}$*) characterizing weights. The estimated SWF weights are characterized by the vector* $(\omega_y)_{y \in \mathcal{Y}} = (f_y(\theta))_y$*, where*

[18] We have also extended the representation to allow the SWF weights to depend on the recent history of agents rather than solely on their last productivity level. The estimated results for the SWF are very similar, showing that current productivity is almost a sufficient statistic. For this reason, we directly consider the restriction that the SWF weights only depend on the current productivity level. The results of the extended estimation are reported in Appendix F.

[19] Typically, in the quantitative exercise, there are 10 idiosyncratic states and 4 instruments of the planner.

$\theta \in \mathbb{R}^p$ solves the following system:

$$0 = \sum_{y \in \mathcal{Y}} L_{k,y} f_y(\theta) \text{ for all } k = 1, \ldots, n, \tag{20}$$

$$1 = \sum_{y \in \mathcal{Y}} \pi_y f_y(\theta). \tag{21}$$

As a robustness check, we also consider a non-parametric estimation of the weights. In this case, the system (20)–(21) is underdetermined and the weights are chosen as the solution of (20)–(21) with the lowest variance across productivity levels. Both solutions imply weights that are quantitatively very similar, which we see as positive for the identification strategy. See Appendix F for definitions and results in the non-parametric case.

The IWF weights. The identification of the IWF weights is of higher dimensionality and requires to compute the $|\mathcal{Y}| \times |\mathcal{Y}|$ parameters $\hat{\omega}_{\tilde{y}y}$. As explained in Section 3.3 for the simple setup, we choose the weights $\tilde{\omega}$ that are the closest to the self-interested ones, while being consistent with the SWF weights. The following definition formalizes it.

Definition 3 *We consider as given SWF weights $(\omega_y)_y$ and policy weights $(\omega_{P,y})_y$. The estimated IWF weights are given by the matrix $(\tilde{\omega}_{\tilde{y}y})_{\tilde{y},y}$ that solves the following program:*

$$(\tilde{\omega}_{\tilde{y}y})_{\tilde{y},y} = argmin_{(\hat{\omega}_{\tilde{y}y})_{\tilde{y},y}} \sum_{(y,\tilde{y}) \in \mathcal{Y}^{\infty 2}} \pi_{\tilde{y}} \left(\hat{\omega}_{\tilde{y}y} - \frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y} \right)^2,$$

$$s.t.\ \omega_y = \sum_{\tilde{y} \in \mathcal{Y}^\infty} \pi_{\tilde{y}} \omega_{P,\tilde{y}} \hat{\omega}_{\tilde{y}y}.$$

The solution to this program is for $(y, \tilde{y}) \in \mathcal{Y}^{\infty 2}$:

$$\hat{\omega}_{\tilde{y}y} = \frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y} + \frac{\omega_{P,\tilde{y}}}{\sum_{\tilde{y} \in \mathcal{Y}^\infty} \pi_{\tilde{y}} (\omega_{P,\tilde{y}})^2} (\omega_y - 1). \tag{22}$$

The proof for the derivation of the weight expression can be found in Appendix A.6. The IWF weights in equation (22) are quite straightforward to interpret. They are separable in two terms, which are the self-interested weights and the distance of the SWF weights to utilitarian ones, weighted by a factor proportional to the political weight $\omega_{P,\tilde{y}}$. Formally:

$$\tilde{\omega}_{\tilde{y}y} = \underbrace{\frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y}}_{=\text{self-seeking weights}} + \frac{\omega_{P,\tilde{y}}}{\sum_{\tilde{y} \in \mathcal{Y}^\infty} \pi_{\tilde{y}} (\omega_{P,\tilde{y}})^2} \times \underbrace{(\omega_y - 1).}_{=\text{distance to utilitarian SWF}} \tag{23}$$

This simple decomposition mostly comes from choosing $(\frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y})_{y,\tilde{y}}$ for the benchmark weights to which the distance should be computed. Indeed, it implies that the IWF weights reduce to self-interested ones ($\tilde{\omega}_{\tilde{y}y} = 0$ if $y \neq \tilde{y}$) for utilitarian SWF weights ($w_y = 1$ for all y). It can be shown that they are the only benchmark for which this property holds.

To push the interpretation further, we consider the quantity $\pi_{\tilde{y}} \omega_{P,\tilde{y}} \hat{\omega}_{\tilde{y}y}$, which is the measure of how much the perception of agents $\tilde{y}$ contributes to the SWF weight ω_y of agent y. We have: $\pi_{\tilde{y}} \omega_{P,\tilde{y}} \hat{\omega}_{\tilde{y}y} = 1_{y=\tilde{y}} + \frac{\pi_{\tilde{y}} (\omega_{P,\tilde{y}})^2}{\sum_{\tilde{y} \in \mathcal{Y}^\infty} \pi_{\tilde{y}} (\omega_{P,\tilde{y}})^2} (\omega_y - 1)$. It includes a visible self-interested component: $1_{y=\tilde{y}}$, showing that along this dimension, the perception of agents $\tilde{y}$ to the SWF matter only when they perceive themselves. The second component can be perceived as an altruistic dimension when it is positive, or a spiteful component when it is negative. All agents $\tilde{y}$ will perceive agents y proportionally to the distance of the SWF weights of these agents to utilitarian. The loading factor put by agents $\tilde{y}$ is proportional to $\pi_{\tilde{y}}(\omega_{P,\tilde{y}})^2$, which is increasing in the population share and political weights of agents $\tilde{y}$.

5 The general model and the Ramsey program

We now construct the macroeconomic model allowing for estimation of the SWF and IWFs. We consider a mass 1 of ex-ante identical agents that is affected by a productivity risk denoted by y – the risk structure is the same as in Section 4. We further assume two goods in the economy: a final consumption good, whose consumption is denoted by c and labor, whose supply is denoted by l. The rest of the specification involves: the planner's fiscal structure in Section 5.1 and the households' program and the competitive equilibrium in Section 5.2. The corresponding Ramsey program and its FOCs are described in Section 5.3. We finally discuss the identification of weights in Section 5.5.

5.1 Production and government

Production. In any period t, a production technology with constant returns to scale transforms capital K_{t-1} and labor L_t into $F(K_{t-1}, L_t,)$ units of output. The production function is smooth in K and L, satisfies the standard Inada conditions, and exhibits constant-to-scale returns. This formulation allows for capital depreciation, which is subsumed by the production function F. Labor L_t is the total labor supply measured in efficient units. The good is produced by a unique profit-maximizing representative firm. We denote by $\tilde{w}_t$ the real before-tax wage rate in period t and by $\tilde{r}_t$ the real before-tax rental rate of capital in period t. Profit maximization yields in each period $t \geq 1$:

$$\tilde{r}_t = F_K(K_{t-1}, L_t) \quad \text{and} \quad \tilde{w}_t = F_L(K_{t-1}, L_t). \tag{24}$$

Government. A benevolent government must finance a path of public spending, (G_t), using several instruments. First, the government can levy one-period public debt B_t, assumed to be default-free. As there is no aggregate risk, public debt and capital are perfectly substitute and they payoff the same pre-tax interest rate $\tilde{r}_t$. Second, the government can raise a number of distortionary taxes, which concern consumption, labor income, and capital revenues. Consumption and capital taxes are linear and are denoted by τ_t^c, and τ_t^K at date t. Regarding the tax on labor income, note that the pre-tax labor income of an agent with productivity y and labor supply l is $\tilde{w}yl$. The associated labor income tax, denoted by $\mathcal{T}_t(\tilde{w}yl)$, is assumed to be non-linear and possibly time-varying, as in Heathcote et al. (2017) (henceforth, HSV):

$$\mathcal{T}_t(\tilde{w}yl) := \tilde{w}yl - \kappa_t(\tilde{w}yl)^{1-\tau_t}, \tag{25}$$

where κ captures the level of labor taxation and τ the progressivity. Both parameters will be planner's instruments. When $\tau_t = 0$, labor tax is linear with a rate $1 - \kappa_t$. Oppositely, the case $\tau_t = 1$ corresponds to full income redistribution, where all agents earn the same post-tax income κ_t. Functional form (25), combined with a linear capital tax, allows one to realistically reproduce the actual US system and its progressivity.[20]

These three taxes imply a total governmental revenue equal to $\tau_t^c C_t + \int_i \mathcal{T}_t(\tilde{w}_t y_{i,t} l_{i,t})\ell(di) + \tau_t^K \tilde{r}_t(K_{t-1} + B_{t-1})$, where C_t is the aggregate consumption, and $A_{t-1} := K_{t-1} + B_{t-1}$ is the aggregate savings in period $t-1$ and $\tilde{r}_t A_{t-1}$ the capital revenues in period t.

With these elements, the governmental budget constraint can be written as follows:

$$G + (1 + \tilde{r}_t)B_{t-1} = \tau_t^c C_t + \int_i \mathcal{T}_t(\tilde{w}_t y_{i,t} l_{i,t})\ell(di) + \tau_t^K \tilde{r}_t A_{t-1} + B_t. \tag{26}$$

We define the post-tax rates r_t and w_t, as follows:

$$r_t := (1 - \tau_t^K)\tilde{r}_t, \quad w_t := \kappa_t(\tilde{w}_t)^{1-\tau_t}. \tag{27}$$

Using the property of constant-return-to-scale for F and the definition of post-tax rates (27), the governmental budget constraint can be written as:

$$G + (1 + r_t)B_{t-1} + w_t \int_i (y_{i,t} l_{i,t})^{1-\tau_t} \ell(di) + r_t K_{t-1} = \tau_t^c C_t + F(K_{t-1}, L_t) + B_t. \tag{28}$$

5.2 Households program

Period utility. We specify the period utility function U of agents. It is defined over private consumption c and labor supply l, and is assumed to be separable. Formally:

$$U(c, l) := u(c) - v(l). \tag{29}$$

The function $u : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable, strictly increasing, and strictly concave, with $u'(0) = \infty$, while $v : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable, strictly increasing, and strictly convex, with $v'(0) = 0$.[21]

Agents' program. Agents' resources consist of labor income and saving payoffs. The post-tax labor income of an agent with productivity $y_{i,t}$ and supplying labor effort $l_{i,t}$ amounts to $\tilde{w}_t y_{i,t} l_{i,t} - \mathcal{T}_t(\tilde{w}_t y_{i,t} l_{i,t}) = w_t(y_{i,t} l_{i,t})^{1-\tau_t}$. Since public debt and capital shares are perfect substitutes, savings payoffs are equal to $(1 + r_t)a_{i,t-1}$, where $a_{i,t-1}$ is the end-of-period-$t-1$

[20] The literature uses either the combination of a linear tax and of a lump-sum transfer (e.g., Dyrda and Pedroni, 2022, Açikgöz et al., 2022) or the HSV structure (see Ferriere and Navarro, 2023). Heathcote and Tsujiyama (2021) show that the HSV structure is quantitatively more relevant.

[21] Without loss of generality we can assume that U is positive for all choices actually made by agents. We can indeed shift u or v by a harmless constant. Note that this constant has no effect on our estimation of the SWF as the strategy of Definition 2 only involves Ramsey FOCs: marginal utilities matter, but utilities in level do not. See Algorithm below.

saving of agent i. Agents use these resources to save and to consume. Formally:

$$\max_{\{c_{i,t},l_{i,t},a_{i,t}\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(c_{i,t}) - v(l_{i,t})\right), \tag{30}$$

$$(1+\tau_t^c)c_{i,t} + a_{i,t} \leq w_t(y_{i,t}l_{i,t})^{1-\tau_t} + (1+r_t)a_{i,t-1}, \tag{31}$$

$$a_{i,t} \geq -\overline{a}, c_{i,t} > 0, l_{i,t} > 0, \tag{32}$$

where $\mathbb{E}_0$ an expectation operator (with respect to idiosyncratic risk), and where the initial state $(y_{i,0}, a_{i,-1})$ is given.

At date 0, agents decide their consumption $(c_{i,t})_{t\geq 0}$, their labor supply $(l_{i,t})_{t\geq 0}$, and their saving plans $(a_{i,t})_{t\geq 0}$ that maximize their intertemporal utility of equation (30), subject to a budget constraint (31) and a previous borrowing limit (32), while prices are assumed to be exogenous. These decisions are functions of the initial endowment $a_{i,-1}$, and the history of idiosyncratic shocks y_i^t. However, to simplify notation, instead of writing the agents' optimal decision as a function of these variables (as was done in Section 4), we simply denote it with the subscripts i and t. For instance for a generic variable X, instead of using the dependence in the history y_i^t, we simply write it $X_{i,t}$. Similarly, instead of summing over all histories in period t, we simply sum over all agents in a given period: $\int_i X_{i,t}\ell(di)$, where ℓ is the distribution of agents on the population interval J.

The FOCs associated with the agents' program (30)–(32) can be written as follows:

$$u'(c_{i,t}) = \beta\mathbb{E}_t\left[(1+r_{t+1})\frac{1+\tau_t^c}{1+\tau_{t+1}^c}u'(c_{i,t+1})\right] + \nu_{i,t}, \tag{33}$$

$$v'(l_{i,t}) = \frac{1-\tau_t}{1+\tau_t^c}w_t y_{i,t}(y_{i,t}l_{i,t})^{-\tau_t}u'(c_{i,t}), \tag{34}$$

where the quantity $\beta^t\nu_{i,t}$ denotes the Lagrange multiplier on agent i's credit constraint at t.

Market clearing and resources constraints. The clearing conditions for capital and labor markets can be written as follows:

$$\int_i a_{i,t}\ell(di) = A_t = B_t + K_t, \quad \int_i y_{i,t}l_{i,t}\ell\,(di) = L_t. \tag{35}$$

Equilibrium definition. We provide a formal definition in Appendix B. Intuitively, for a given fiscal policy $(\tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t)_t$, the competitive equilibrium is a collection of individual decisions $(c_{i,t}, l_{i,t}, a_{i,t}, \nu_{i,t})_{t,i}$, of aggregate quantities $(K_t, L_t, Y_t)_t$, and prices $(w_t, r_t, \tilde{w}_t, \tilde{r}_t)_t$ that are consistent with the agents' optimization program (30)–(32), the clearing equation (35) of financial and labor markets, as well as the definition of pre- and post-tax factor prices (24) and (27). The competitive equilibrium is at the steady state when all quantities are time-invariant.

5.3 The Ramsey problem and the identification of weights

We follow the construction of Proposition 1 for the SWF. We also require Assumption A to hold for identification purposes. With our period utility separable in consumption and labor, the period 0 SWF is:

$$SWF_0 := \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \omega_{y_{i,t}} \left(u(c_{i,t}) - v(l_{i,t}) \right) \ell(di) \right], \tag{36}$$

where the weights $(\omega_y)_{y\in\mathcal{Y}}$ solely depend on the current productivity level due to Assumption A.

In the Ramsey program, the planner aims to determine the fiscal policy corresponding to the competitive equilibrium that maximizes aggregate welfare according to the criterion in equation (36), while satisfying the government's budget constraint. A Ramsey equilibrium is a fiscal policy, prices, individual allocations and aggregate quantities solving the Ramsey program. A Ramsey steady state equilibrium is a time-invariant Ramsey equilibrium. Formally, the Ramsey program can be stated as follows.

$$\max_{\left(w_t, r_t, \tilde{w}_t, \tilde{r}_t, \tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t, G_t, K_t, L_t, (c_{i,t}, l_{i,t}, a_{i,t}, \nu_{i,t})_i\right)_{t\geq 0}} W_0, \tag{37}$$

$$G_t + (1+r_t)B_{t-1} + w_t \int_i (y_{i,t} l_{i,t})^{1-\tau_t} \ell(di) + r_t K_{t-1} = \tau_t^c C_t + F(K_{t-1}, L_t, s_t) + B_t, \tag{38}$$

$$\text{for all } i \in \mathcal{I}\text{: } (1+\tau_t^c)c_{i,t} + a_{i,t} = (1+r_t)a_{i,t-1} + w_t (y_{i,t} l_{i,t})^{1-\tau_t}, \tag{39}$$

$$a_{i,t} \geq -\bar{a}, \ \nu_{i,t}(a_{i,t} + \bar{a}) = 0, \ \nu_{i,t} \geq 0, \tag{40}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t \left[(1+r_{t+1}) \frac{1+\tau_t^c}{1+\tau_{t+1}^c} u'(c_{i,t+1}) \right] + \nu_{i,t}, \tag{41}$$

$$v'(l_{i,t}) = \frac{(1-\tau_t)}{1+\tau_t^c} w_t y_{i,t} (y_{i,t} l_{i,t})^{-\tau_t} u'(c_{i,t}), \tag{42}$$

$$K_t + B_t = \int_i a_{i,t} \ell(di), \ L_t = \int_i y_{i,t} l_{i,t} \ell(di), \tag{43}$$

and subject to the definition (24) of pre-tax wage and interest rates $\tilde{w}_t$ and $\tilde{r}_t$, the definition (27) of post-tax rates, the positivity of labor and consumption choices, and initial conditions.

Since the Ramsey program involves selecting a competitive equilibrium, its constraints include the equations characterizing this equilibrium: individual budget constraints (39), individual credit constraints (and related constraints on $\nu_{i,t}$) (40), Euler equations for consumption and labor - (41) and (42) - and market clearing conditions for financial and labor markets (43). Moreover, the fiscal policy selected by the Ramsey equilibrium should also fulfill the governmental budget constraint (38) – which is also a constraint.

The general Ramsey program can be simplified. First, in our setup with a linear tax on capital, a progressive tax on labor, and one-period public debt, the consumption tax is redundant with other fiscal instruments. Second, we can follow Chamley (1986) and express the program using post-tax prices only. Combining the two simplifications implies that the planner's fiscal instruments are: post-tax wage and interest rates W_t and R_t, labor tax progressity and public debt. They need to be chosen such that the governmental budget constraint (28) holds. The other choice variables of the planner also include individual and aggregate allocations that have to be chosen so as to correspond to a competitive equilibrium. This means that individual budget constraints (31), borrowing limits (32), and FOCs (33)–(34) are constraints of the Ramsey program – as well as market clearing conditions (35). The reformulated Ramsey program is formally stated in Proposition 3 of Appendix C.1.

5.4 Interpreting the Ramsey FOCs in the light of public finance

The economic trade-offs faced by the planner can be identified by the FOCs of the Ramsey program, which can be found in Appendix C.2.[22] We here discuss the economic interpretation of the FOCs of the Ramsey program using the concepts of public finance and extending our discussion of Section 3.3.

We focus here on an arbitrary fiscal instrument $(I_t)_t$, which in our context can be the capital tax (or the post-tax instrument R_t), the labor tax level (or the post-tax instrument W_t), or the labor tax progressivity. This analysis thus includes all instruments except public debt, which is discussed below. We consider that the planner raises resources through a variation of the fiscal instrument, which decreases the consumption of all agents. The Lagrangian associated to problem (37) can be written as:

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\mathcal{W}_t + \mu_t \mathcal{B}_t \right).$$

First, $\mathcal{W}_t$ is the "augmented welfare" of date t that includes the pure welfare component $\omega_{y_{i,t}}(u(c_{i,t}) - v(l_{i,t}))$, as well as the general-equilibrium effects implied by individual decisions about savings (i.e., Euler equation) and labor supply (i.e., FOC labor supply).[23] This term depends on the fiscal instrument I_t, and on all consumption levels $(c_{i,t})_i$ of date t. Second, $\mathcal{B}_t$ is the governmental budget constraint at date t and μ_t the associated Lagrange multiplier – we keep the same notation as in the simple case. As shown in Appendix C.2, the quantity $\mathcal{B}_t$ depends on I_t but not on $(c_{i,t})_i$. With this notation, the FOC with respect to I_t can be written as:[24]

$$\mu_t = \int_i \underbrace{\left.\frac{\partial \mathcal{W}_t}{\partial c_{i,t}}\right|_{I_t}}_{=SVL_{i,t}} \underbrace{\frac{-\frac{\partial c_{i,t}}{\partial I_t}}{\frac{\partial \mathcal{R}_t}{\partial I_t} + \frac{1}{\mu_t} \left.\frac{\partial \mathcal{W}_t}{\partial I_t}\right|_{c_{i,t}}}}_{=MVPF^I_{i,t}} \ell(di), \tag{44}$$

or

$$\mu_t = \int_i SVL_{i,t} \times MVPF^I_{i,t} \ell(di). \tag{45}$$

As in the simple case, the planner sets the instrument value to the point where the shadow price of the governmental budget constraint (μ_t) equals its marginal cost, which equals to the sum over all agents of the individual (negative) bangs for the buck of a cut in agents' consumption due to a change in the fiscal instrument. The cost covers all agents as the instrument I_t is a non-specific tax. Similarly to the simple case (11), the individual bangs for the buck of one unit of resources raised through the fiscal instrument I equals the product of SVL, $SVL_{i,t}$ (which is independent of the fiscal instrument and generalizes the notion of social marginal weight), and the MVPF, denoted $MVPF^I_{i,t}$ (which is instrument-specific).

Similarly to the simple model, $SVL_{i,t}$ (which we denote by $\psi_{i,t}$ in Appendix C) quantifies the welfare reduction (increase) associated to a one-unit reduction in the consumption of agents

[22] Solving such a program through a Lagrangian raises a number of technical questions that have been discussed in LeGrand and Ragot (2024) in details.

[23] The consumption tax is redundant and as such does not imply any specific FOC.

[24] $\left.\frac{\partial \mathcal{W}_t}{\partial c_{i,t}}\right|_{I_t}$ is the partial derivative of $\mathcal{W}_t$ with respect to $c_{i,t}$ while keeping I_t constant; $\left.\frac{\partial \mathcal{W}_t}{\partial I_t}\right|_{c_{i,t}}$ is the partial derivative of $\mathcal{W}_t$ with respect to I_t while keeping all $c_{i,t}$ constant.

i, regardless of the fiscal instrument causing this reduction. As explained above, this welfare reduction includes the endogenous effects on labor supply and savings generated by the variation in consumption. The total welfare effect thus includes three terms:

$$SVL_{i,t} = \underbrace{\omega_{i,t}u'(c_{i,t})}_{=\text{direct effect}} - \underbrace{(\lambda_{c,i,t} - R_t\lambda_{c,i,t-1})u''(c_{i,t})}_{=\text{externality on savings incentives}} + \underbrace{\lambda_{l,i,t}(1-\tau_t)W_t(y_{i,t})^{1-\tau_t}(l_{i,t})^{-\tau_t}u''(c_{i,t})}_{=\text{externality on labor supply incentives}}, \tag{46}$$

where $\beta^t\lambda_{c,i,t}$ and $\beta^t\lambda_{l,i,t}$ are the Lagrange multipliers on agent i's Euler equation and labor supply FOC, respectively. The first term, $\omega_{i,t}u'(c_{i,t})$, reflects the direct welfare effect of a consumption variation. This is identical to the social marginal welfare weight in equation (11). However, in our setting, welfare is also affected by the changes in savings and labor supply induced by the variation in consumption. The welfare impact due to savings channels through the Euler equation, while the one due to labor supply channels through the FOC on labor supply. This explains why the indirect welfare effects of savings and labor supply are proportional to the Lagrange multipliers. The notion of SVL thus generalizes the notion of MSWW to endogenous labor and savings choices.[25]

The MVPF measures the variation (here, a decrease) in consumption implied by a one resource unit taken from agent i by a marginal variation of the fiscal instrument I_t, including all fiscal externalities implied by the instrument. In the term $MVPF^I_{i,t}$ of equation (44), in the absence of direct pecuniary externality of I_t, $\frac{\partial \mathcal{B}_t}{\partial I_t}$ measures the direct effect of the fiscal instrument on planner's resources and is equal to the instrument fiscal base (i.e., all payoffs of interest-bearing assets for the capital tax or the labor supply for the labor tax level). Conversely, $\frac{1}{\mu_t}\frac{\partial \mathcal{W}_t}{\partial I_t}\Big|_{c_{i,t}}$ is the fiscal externality of I_t that channels through the modification of the savings incentives and the Euler equation. This fiscal externality reflects that the instrument I_t is distortionary.

Note that if the planner would have access to a standard aggregate lump-sum transfer T, the MVPF associated to that tax instrument would simply be 1. Indeed, the lump-sum transfer involves no externality and is a flow of resources from the transfer to agents', such that: $\frac{\partial \mathcal{W}_t}{\partial T_t}\Big|_{c_{i,t}} = 0$ and $-\frac{\partial c_{i,t}}{\partial T_t} = \frac{\partial \mathcal{R}_t}{\partial T_t}$. Therefore, the planner would set T_t such that $\mu_t = \int_i SVL_{i,t}\ell(di)$: the marginal cost for governmental budget equals the marginal benefit for all agents. Should the planner would further have access to individual-specific lump-sum transfers T^i, each would be set such that would $\mu_t = SVL_{i,t}$. As explained in LeGrand and Ragot (2024), the difference $SVL_{i,t} - \mu_t$ can thus be thought as capturing the cost of distortionary fiscal instruments for the planner, concerning agent i.

Public debt is the only fiscal instrument for which a FOC similar to equation (45) does not hold. Indeed, public debt at date t affects the contemporaneous governmental budget constraint, $\mathcal{B}_t$ (due to debt issuance), and the one of the next date, $\mathcal{B}_{t+1}$ (due to debt repayment). Furthermore, public debt has no direct impact on households welfare: $\frac{\partial \mathcal{W}_t}{\partial B_s} = 0$. The FOC

[25] This SVL is similar to the quantity $\hat{g}$ defined in Ferey et al. (2024) and that they describe as "the social marginal welfare weights augmented with the fiscal impact of income effects" and which represent "the full social value of marginally increasing the disposable income of [an] individual".

related to debt can be written as $\mu_t \frac{\partial \mathcal{B}_t}{\partial B_t} = \beta \mu_{t+1}(-\frac{\partial \mathcal{B}_{t+1}}{\partial B_t})$: relaxing the budget constraint today comes at the cost of tightening it tomorrow. With the expression of the governmental budget constraint, the public debt FOC becomes $\mu_t = \beta \mu_{t+1}(1 + r_{t+1})$. This FOC is an Euler-like equation, reflecting that the planner uses public debt to smooth out the cost of resources across time.

5.5 Identification of weights

Our estimation involves identifying the SWF weights such that a Ramsey planner endowed with this SWF optimally selects the observed fiscal system and aggregate allocation of a given country. More precisely, we follow the identification strategy of Definition 2 to estimate the social weights $(\omega_y)_{y \in \mathcal{Y}}$ of the SWF SWF_0.

Fiscal policy is composed of five instruments $(\tau^K, \tau^c, B, \kappa, \tau)$, but these five instruments actually impose only two constraints on social weights. Indeed, consumption taxes τ^c are redundant, as explained above (see Appendix C.1 for the details). Second, the public debt FOC, provided in equation (79) of Appendix C.2, imposes a steady-state value on the before-tax real interest rate $1 + \tilde{r} = 1/\beta$, but does not restrict the social weights. Therefore, this means that the instruments $(\tau^K, \tau^c, \kappa, \tau)$ actually imply three FOCs. One of them is used to pin down the Lagrange multiplier of the governmental budget constraint μ_t. The two remaining FOCs imply the two linear constraints on the SWF weights.

The identification strategy of Definition 2 can thus be readily applied with three constraints: the two linear constraints coming from the Ramsey FOCs and the normalization constraint. We thus consider a parametric estimation, with three degrees of freedom that will be exactly identified by the three constraints. We adopt the following functional form, which naturally extend the one in Heathcote and Tsujiyama (2021):

$$\forall y \in \mathcal{Y}, \ \log \omega_y := \bar{\omega}_0 + \bar{\omega}_1 \log(y) + \bar{\omega}_2 (\log(y))^2, \tag{47}$$

where $(\bar{\omega}_i)_{i=1,\dots,3}$ are the three free parameters.

5.6 Solution method and algorithm

The Ramsey problem discussed in Section 5.3 involves a joint distribution across wealth and Lagrange multipliers. This high-dimensional object raises a number of difficulties for the resolution of the Ramsey program. We rely on the truncation method, that has already been used in recent papers (LeGrand and Ragot, 2022b, 2023, 2024). We here improve on previous work to allow for the estimation of the SWF with a utility function separable in consumption and labor, instead of the GHH case considered in previous papers.

The basic idea is to construct a consistent finite state space representation of the model and use it to compute the FOCs of the Ramsey planner. We then use the inverse optimal approach to estimate the SWF.

More precisely, the solution method is based on the following steps.

1. We simulate the heterogeneous-agent model at the steady state, with realistic values for the fiscal instruments and income and wealth inequalities. Standard solution techniques

provide the steady-state distribution of wealth $\Lambda(a, y)$ for any idiosyncratic state $y \in \mathcal{Y}$ and asset holding a, as well as the policy rules for wealth, consumption, and labor supply denoted by $g_a(a, y)$, $g_c(a, y)$ and $g_l(a, y)$, respectively.

2. We consider a given finite set $\mathcal{H}$ of histories for which the transition matrix is a Markov matrix. The most intuitive set of histories is composed of all histories of a given length N. If there are Y idiosyncratic states, there will be Y^N truncated histories.

3. We consider a so-called truncated history $y^N := \{y_1, \ldots, y_N\}$ in the set $\mathcal{H}$, which corresponds to agents experiencing y^N over the last N periods. Using the distribution $\Lambda(a, y_1)$ and the policy rules, we can compute the distribution of wealth, denoted $\tilde{\Lambda}(a, y^N)$, for any truncated history y^N and asset holding a.

4. Using the distribution $\tilde{\Lambda}$, we can aggregate key individual quantities and equations to express them in terms of truncated histories. For instance, the size of truncated history y^N (i.e., the measure of agents with recent history y^N) is $S_{y^N} = \int_0^\infty \tilde{\Lambda}(da, y^N)$, or the per-capita consumption $c_{y^N} = \int_0^{+\infty} g_c(a, y^N)\tilde{\Lambda}(da, y^N)/S_{y^N}$. Finally, average marginal utility is $\int_0^{+\infty} u'(g_c(a, y^N))\tilde{\Lambda}(da, y^N)/S_{y^N} := \xi^u_{y^N} u'(c_{y^N})$, where $\xi^u_{y^N}$ captures both the convexity of the marginal utility and the heterogeneity in the wealth distribution of agents having the same history y^N for the last N periods. The aggregation process thus generates Y^N budget constraints, Euler equations, and labor supply choices (see equations (92)–(94) in Appendix D.1). This defines the *truncated model.*

5. We can compute the FOCs of the planner in the truncated model (see Appendix D.2).

6. We derive the two linear constraints on the SWF weights from the FOCs of the planner. We use them as inputs in (20) for the identification strategy of Definition 2.

7. We consider the following functional form for the weights: $\omega_y := e^{\bar{\omega}_0 + \bar{\omega}_1 \log(y) + \bar{\omega}_2 (\log(y))^2}$ with $(\bar{\omega}_i)_{i=1,\ldots,3}$ being the parameters. We then apply the identification strategy of Definition 2.

8. Using some measure of political weights $\omega_{P,y}$, we determine the IWFs using the expression (22) of the identification strategy of Definition 3.

The detailed derivations of these steps is performed in Appendix D. We consider 10 idiosyncratic states and $N = 5$ as a benchmark, and thus 10^5 possible histories. The estimation process takes less than 3 minutes, and we have checked that the results are robust to an increase in N. We now provide the quantitative investigation, and further discuss the choice of the measures of the political weights in Section 3.4.

6 Quantitative investigation

We first provide the calibrations reproducing the tax system and the wealth distribution in both the US and France for the period 1995-2007. We then estimate the SWFs in both countries.

6.1 Calibration

The US calibration

The estimation parameters are gathered in Table 3, and we detail below our calibration strategy.

Preference parameters. The period is a quarter. The discount factor is set to $\beta = 0.992$ to match a realistic capital-to-output ratio. The period utility functions are $u(c) = \dfrac{c^{1-\gamma} - 1}{1-\gamma}$ and $v(l) = \dfrac{1}{\chi}\dfrac{l^{\frac{1}{\varphi}+1}}{\frac{1}{\varphi}+1}$. We set the inverse of intertemporal elasticity to $\gamma = 1.8$ to match a realistic wealth inequality for the targeted capital-to-output ratio. Furthermore, the Frisch elasticity for labor is set to $\varphi = 0.5$, which is recommended by Chetty et al. (2011). We set the labor-scaling parameter to $\chi = 0.0477$, which implies normalizing the aggregate labor supply to 0.3.

Technology. The production function is of the Cobb-Douglas form and subsumes capital depreciation: $F(K, L, s) = sK^{\alpha}L^{1-\alpha} - \delta K$. The capital share is set to the standard value, $\alpha = 36\%$, while the depreciation rate is set to $\delta = 2.5\%$.

Idiosyncratic labor risk. Various estimations of the idiosyncratic process can be found in the literature. The productivity follows an AR(1) process: $\log y_t = \rho_y \log y_{t-1} + \varepsilon_t^y$, with $\varepsilon_t^y \sim_{\text{IID}} \mathcal{N}(0, \sigma_y^2)$. The calibration features an persistence $\rho_y = 0.99$ and a standard deviation $\sigma_y = 0.0995$, which is close to the estimates of Krueger et al. (2018). We discretize this AR(1) process using the Tauchen (1986) procedure, with 10 states. This calibration implies a Gini index of post-tax and transfers of 0.40, as in Table 1.

Taxes and government budget constraint. Fiscal parameters are calibrated based on the computations by Trabandt and Uhlig (2011) reported in Table 1, with the exception of the progressivity of the labor tax, which we computed ourselves and reported in Table 2. We recall that their estimations for the US in the period 1995-2007 yield a capital tax of $\tau^K = 36\%$ and a consumption tax of $\tau^c = 5\%$. In our estimation for the progressivity parameter, we obtain $\tau = 0.16$, which is close to the estimates in the literature (see Section 2 for the details of our estimation).

Finally, we estimate the parameter κ such that it matches the public-spending-to-GDP ratio of 15%. We obtain a value of $\kappa = 0.85$, which is close to the estimates of Ferriere and Navarro (2023). With this fiscal system, the model generates a public-debt-to-GDP ratio equal to 63%, which corresponds to the value reported in Table 1.

Additionally, the model performs well in replicating the ratios of consumption over GDP and investment over GDP. The model predicts a consumption-to-GDP ratio of 58%, very close to its empirical counterpart of 60% for the period 1995-2007. The investment-to-GDP ratio generated by the model amounts to 27%, close to the empirical value of 25%. Finally, regarding inequalities, the model generates a Gini index for post-tax income equal to 0.40, identical to its empirical counterpart in Table 1. The Gini index for wealth is found to be 0.78, very close to its empirical value of 0.77 in Table 1.

		US		France	
Parameter	Description	Value	Target or ref.	Value	Target or ref.
Preference parameters					
β	discount factor	0.992	$K/Y = 2.7$	0.996	$K/Y = 3.1$
u	utility function	$\cdot$	$\gamma = 1.8$	$\cdot$	$\gamma = 1.8$
φ	Frisch elasticity	0.5	Chetty et al. (2011)	0.5	Chetty et al. (2011)
χ	hours worked	0.33	Penn World Table	0.29	Penn World Table
α	capital share	36%	Profit Share, NIPA	36%	Profit Share, INSEE
δ	depreciation rate	2.5%	Chetty et al. (2011)	2.5%	Own calc., INSEE
Productivity parameters					
σ^y	std. err. prod.	0.10	Gini for income	0.06	Fonseca et al. (2023)
ρ^y	autocorr. prod.	0.99	Gini for income	0.99	Fonseca et al. (2023)

Table 3: Parameter values.

We gather the model implications in Table 4. These implications show that our tax system provides a good approximation of the income and wealth distribution in the US, and hence of the redistributive effects of the US tax system. This confirms the results of Heathcote et al. (2017) and Dyrda and Pedroni (2022).

French calibration

The calibration for France shares a number of similarities with the one for the US. We use the same period and the same functional forms. For the sake of clarity, we mimic the structure of the US calibration, even though our presentation is more streamlined. The calibration parameters can be found, as those for the US, in Table 3.

Preference parameters. The discount factor is set to $\beta = 0.996$ and the Frisch elasticity for the labor supply is still equal to $\varphi = 0.5$. We fix the scaling parameter to $\chi = 0.0228$, which implies an aggregate labor supply normalized to 0.3. It happens that the same risk-aversion parameter $\gamma = 1.8$ is consistent with French statistics.

Technology and TFP shock. We keep the same production function: $F(K, L, s) = sK^{\alpha}L^{1-\alpha} - \delta K$, with the same parameter values: $\alpha = 36\%$ and $\delta = 2.5\%$.[26]

Idiosyncratic risk. The AR(1) productivity process is calibrated using $\rho_y = 0.99$ and $\sigma_y = 0.0646$. These values are in line with the estimates of Fonseca et al. (2023). As for the US calibration, we discretize this process with 10 states.

Taxes and government budget constraint. We use the values summarized in Table 1 for the French taxes, except for the labor tax that is progressive. We consider a capital tax of $\tau^K = 35\%$, a progressivity parameter of $\tau = 0.23$, and a consumption tax of $\tau^c = 18\%$. This tax system has realistic implications for the model. In terms of public finance, we use $\kappa = 0.728$ to

[26]We are keeping the same values as in the United States to emphasize that the differences in the SWFs are due to differences in the fiscal systems and not due to different production functions.

Parameter	Description	US Model	US Data	France Model	France Data
Public finance aspects					
B/Y	Public debt (%GDP)	63%	63%	60%	60%
G/Y	Public spending (%GDP)	15%	15%	25%	24%
	Total tax revenues (%GDP)	16%	26%	25%	40%
Aggregate quantities					
C/Y	Aggregate consumption (%GDP)	58%	60%	44%	45%
I/Y	Aggregate investment (%GDP)	27%	25%	31%	31%
Inequality measures					
	Gini for post-tax income	40%	40%	28%	28%
	Gini for wealth	78%	77%	68%	68%

Table 4: Model implications for key variables. Empirical values are discussed in Section 2 and summarized in Table 1.

match the empirical public-spending-to-GDP of 24%. This implies a public-debt-to-GDP ratio of 60%, which matches the value of Table 1. Regarding private consumption and investment, the model generates aggregate private consumption equal to 44% of GDP, which is close to the empirical counterpart of 45% estimated by Trabandt and Uhlig (2011) for the period 1995-2007, while investment amounts to 31% of GDP, equal to its empirical counterpart. Finally, in terms of inequalities, the model implies a Gini index for post-tax income of 0.28 and a Gini index for wealth of 0.68. These two Gini values match their empirical counterparts of Table 1. Again, this confirms that the tax system is empirically relevant.

6.2 Estimation of the SWFs

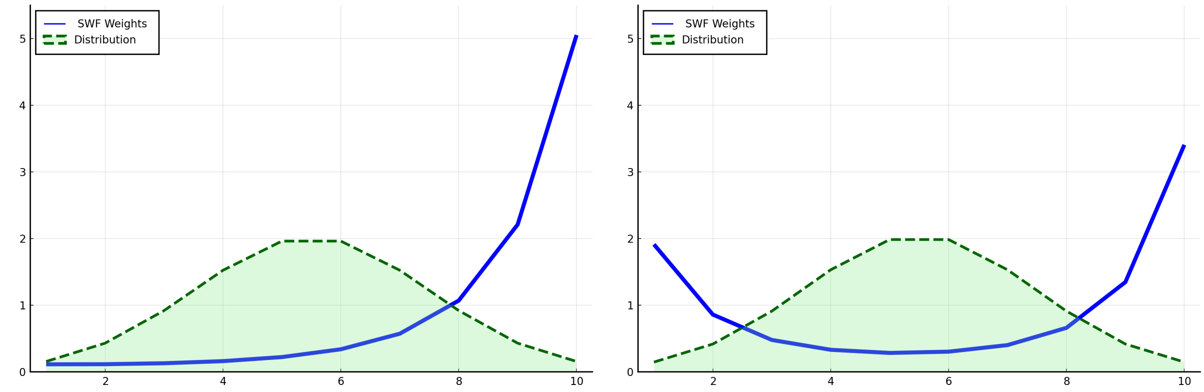
The estimation procedure follows the algorithm presented in Section 5.3 and the algebra of Appendix D. For the simulations below, we consider a truncation length of $N = 5$, although the main characteristic of the results do not change when we consider longer truncation lengths. As there are 10 idiosyncratic productivity levels, the number of truncated histories amounts to $N^{tot} = 10^5 = 100000$.

As discussed in Section 5.5, the weights are obtained such that the FOCs of the planner are exactly identified. We apply Algorithm of Section 5.6 and we obtain the following parametric function for the US and France, respectively:

$$\log \omega(y)^{us} = -0.25 + 1.06 \log(y) + 0.22(\log(y))^2,$$
$$\log \omega(y)^{fr} = -0.51 + 0.62 \log(y) + 1.44(\log(y))^2.$$

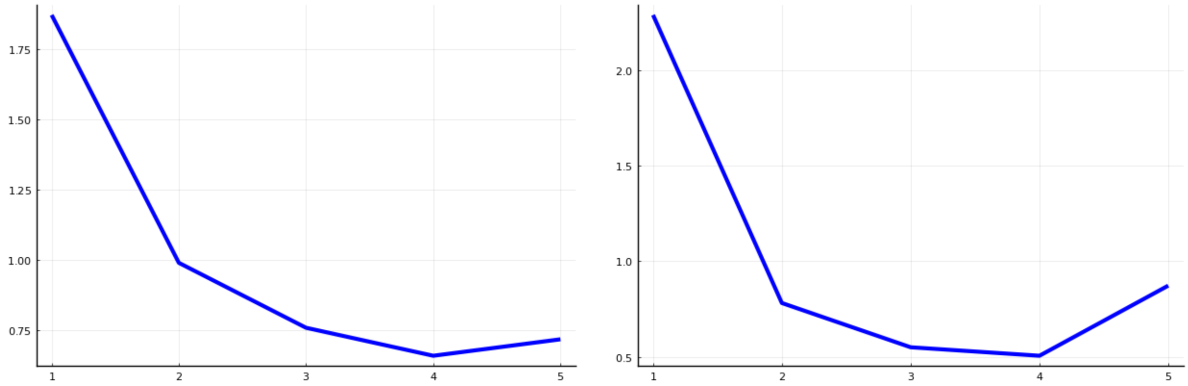
In Figure 2, we plot the weights of the SWF as a function of the 10 productivity indices of agents. We observe that in the US, the period weights increase with productivity level, whereas for France they exhibit a U-shaped pattern, assigning higher weights to low-productivity agents compared to those at the top of the productivity distribution.

In the US, agents with the highest weight in the population are those with the highest productivity. In France, low-productivity agents have a higher weight than those with medium



(a) United States (b) France

Figure 2: Parametric period weights as a function of productivity for the US and France.



(a) United States (b) France

Figure 3: Mean SMWWs per quintile as a function of the quintile of labor income in the United States and in France.. See text for comments.

productivity. The high productivity agents have the highest weights.

Implied marginal weights

As discussed in Section 5.4, the public finance literature often considers the SMWWs, which is the product of the social weights by the average marginal utility for each productivity level: $\omega_i \bar{u'}_i$. Although the relevant concept for the planner in our environment is the SVL (see equation (46)), it is useful to represent implied SMWWs, as they have been estimated for the US (e.g., Hendren, 2020). Figure 3 represents the mean SMWWs in each quintile of the labor income, as a function of the labor income quintile. SMWWs are normalized such that they average to 1 across quintiles.

The shape of mean SMWWs is similar in both countries. The SMWWs are decreasing with income quintile, except for the last quintile for which they are increasing. This shows that the same shape for SMWWs can be consistent with very different SWF weights.

Interestingly, the shape for the US is similar to those estimated by Hendren (2020) using fiscal data – to our knowledge, there is no such estimation for France. The SMWWs are found to be decreasing with income quantile, except at the end of distribution, where they slightly increase. Hendren's weights for first quintile have a lower value than ours: They amount to 1.2, while ours are about 1.8. [27] This difference comes from the fact that low productivity agents have a very high marginal utility in our case, as we do not take into account some money transfers, which are captured in Hendren. Indeed, our fiscal system, although relevant for macroeconomics, is very simple compared to the actual transfer scheme in the US.

Despite these differences, we consider that the similarity in the general shape is promising and encouraging. It shows that heterogeneous-agent models can be consistent with the empirical public finance literature.

6.3 Investigating the drivers of the weight differences between the US and France

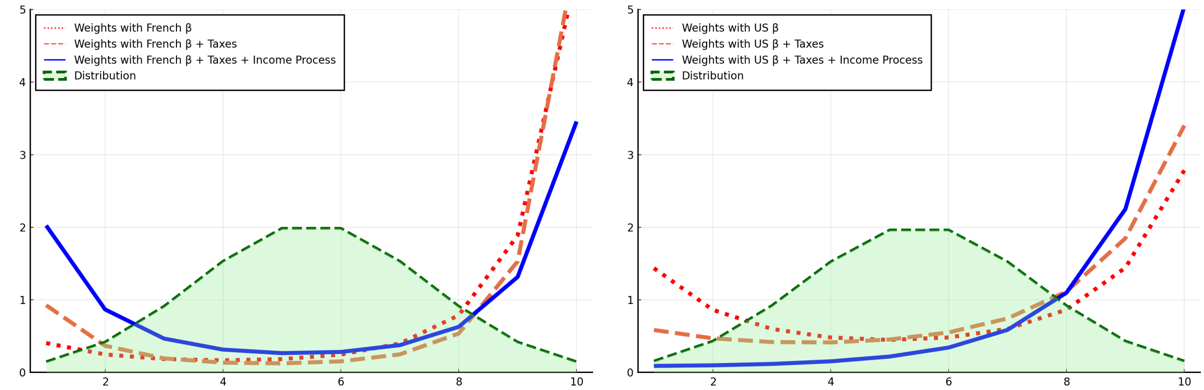
Before interpreting these weights, we use the previous methodology to investigate the drivers behind the differing weights assigned to agents in the United States and France. The objective of this section is to understand why the weights differ so much between France and US. We decompose the differences along the three sources of heterogeneity between the two countries: (i) the discount factor β; (ii) the fiscal system; (iii) the productivity process. Indeed, the calibrations of the two countries only differ only these three lines.

We start with the role of the discount factor. In panel (a) of Figure 4, the red dashed line represents the SWF weights as a function of productivity for the US calibration, except the discount factor which is set to the French value. Compared to the original weights, there is a slight increase in the weights for low productivity agents, but the overall shape remains similar: higher weights are given to agents with higher productivity levels. Similarly, in panel (b), the red dashed line plots the weights for France adopting the US discount factor. We observe that the weights for low-productivity agents in France decrease, while those for high-productivity agents increase. However, the discount factor alone does not fully account for the differences in weights between the two countries. Overall, making agents and the planner more patient (i.e., increasing β) tends to increase the weights of low-productivity agents, and to decrease those of high-productivity agents.

Second, we analyze the impact of fiscal systems. Panel (a) of Figure 4 shows the US weights with the French tax system and the French β (orange dashed line). The weights for lower productivity agents increase at the expense of those for higher productivity agents. Conversely, in panel (b) of Figure 4, we plot the weights when France adopts the US tax system in addition to the US β. The results mirrors those of the US: The weights for low-productivity agents decrease, while those for high-productivity agents increase. This exercise illustrates the role of the fiscal system. The French tax system, characterized by a higher progressivity and a greater inclination to the reduction of inequality, contributes to increase the weights of lower productivity agents at the expense of those of higher productivity ones. The role of the US tax system, which is more Libertarian (as we will discuss below), has an opposite effect.

Finally, to fully uncover the differences in weights, we incorporate the French income process into the US economy in addition to the French β and the French tax system. the resulting weights then exactly replicate the French weights (blue line in panel (a) of Figure 4). This is a mechanical result as in that case the modified US economy has the same calibration the baseline

[27] These different initial values also change the concavity of the SMWWs relationships.



(a) United States (b) France

Figure 4: Change in the period weights for the United States and France due to different preference parameters, tax systems, and income processes.

French economy. We observe the same result for France (panel (b)).

We analyze further the impact for the US of opting for the French tax system – and the other way around. Figure 8 in Appendix E illustrates the role of the fiscal system on SWF weights, utility, labor, and capital income. Adopting the French tax system in the US reduces labor income for high-productivity agents, because it reduces labor supply incentives. Thus, it also reduces reduces the utility of these high-productivity agents. This heavier labor taxation results from lower weights assigned to high-productivity agents by the social planner. Conversely, the progressive tax system boosts the consumption and the utility of low-productivity agents, and is the result for their larger SWF weights. This experiment demonstrates that changes in the tax system can rationalize changes in weights. For the US to increase weights for low-productivity agents and decrease weights for high-productivity agents, adopting a more progressive labor tax is effective. The relatively low labor tax in the US favors high-income/high-productivity agents.

6.4 A world where the US have the French SWF

We now compute the US fiscal system that makes SWF weights as close as possible to those of France. The goal is to find a fiscal system in the US where the distance between the weights in the modified US economy and the benchmark France economy is minimized. This exercise aims to understand the role of social preferences in shaping the tax system, distinct from the influence of technology and individual preferences.

We conduct the experiment as follows. We start from the calibration of the US: we consider the preference parameters, the production function, and the productivity process of the US. Independently of the fiscal scheme, this sets the steady-state value of the capital-to-output ratio. We then iterate over the capital tax rate and the progressivity of the labor tax to minimize the distance between the corresponding SWF weights and the French ones. We keep adjusting the parameter κ, driving the labor tax level, to keep the government spending-to-output ratio equal to its US counterpart. This means that for any fiscal policy, the main macroeconomic ratios (capital-to-output, investment-to-output, aggregate consumption to output and public spending to output) in the fictive economy are identical to those in the US.

There is subtlety in the computation of the distance between the SWF weights. On the one

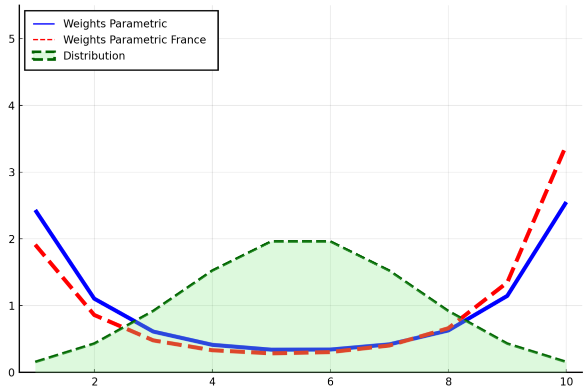


Figure 5: SWF weights for France (red dashed line) and for the United States with the tax system that minimizes the distance to the French weights (blue line). The x-axis corresponds the 10 productivity levels in France.

hand, the benchmark SWF weights of France correspond to the French productivity levels. On the other hand, the weights we calculate in the fictive economy correspond to the US productivity levels, as we are considering the productivity process of the US. To compute the distance, we define the weights of the two economies for the same productivity levels. We therefore interpolate the weights to obtain their values for both the US and the French productivity levels. The objective we minimize is thus the Euclidean distance between the SWF weights computed for French and US productivity levels.

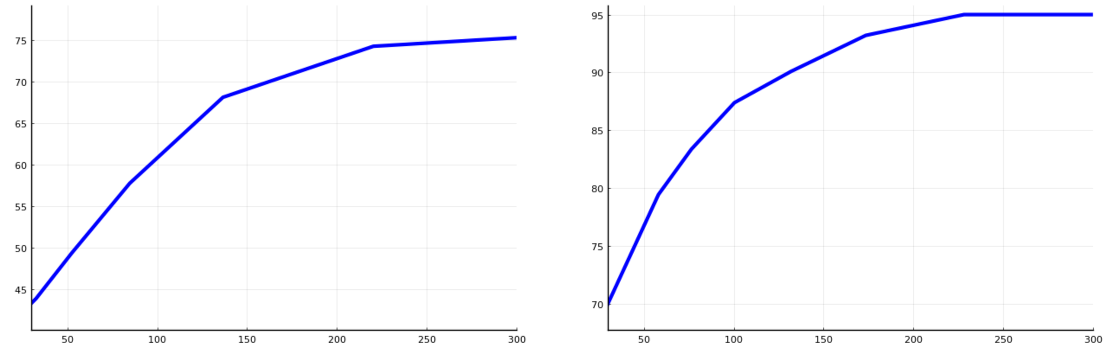
The minimization yields a new fiscal system that corresponds to the "core" US economy with the French SWF. We refer to this economy as the *US with French SWF*. Figure 5 plots the weights of the *US with French SWF* and the weights of France as a function of the French productivity levels. As can be seen, the minimization procedure is successful in finding a fiscal system that allows the SWF weights of the two economies to be quite close to each other.

We report in Table 5 the values of the new fiscal system in the *US with French SWF* economy (row *US with French SWF*). For the sake of comparison, we also report the fiscal system of the (baseline) US and French economies (rows *US* and *France*).As can be seen from the Gini values, the distribution of income and wealth of the US with French SWF is now much closer to their French counterparts, and therefore less unequal than in the US.

	Public debt (%GDP)	τ_k (%)	τ (%)	κ (%)	Gini post-tax income	Gini wealth
US	63	36	16	85	40	78
France	60	35	23	73	28	68
US with French SWF	299	9	57	71	27	63

Table 5: Comparison between the benchmark economies and the US economy with the French SWF.

This reduction in equality mostly comes from the higher weight of low-productivity/low-income agents. This higher weight translates into a much higher progressivity. The progressivity indeed increases from 16% to 57%. We recall that the other "core" parameters, as well as the main



a) United States b) France

Figure 6: Participation rates (in percent) as a function of annual income (expressed as a percentage of average income). US data refer to the 2008 presidential election. Data for France are for the 2007 presidential election. The 100 on the x-axis is the average income in each country.

macro ratios (e.g., consumption-to-GDP, government spending-to-GDP, investment-to-GDP) remain the same as in the benchmark US economy. In particular, the pre-tax interest rate is kept at its optimal value, which is the inverse of the discount factor. Because of the labor taxation, the labor supply falls, which means that the capital also falls to keep the capital-to-labor ratio constant. However, agents still have the US productivity levels, which makes them overall save more than in France. This requires an increase in public debt to absorb the excess savings.

This higher progressivity is detrimental to high-productivity agents. However, they have a quite a large weight in the French SWF. To partly offset for those agents the large progressivity increase, the capital tax is lowered. This lower capital tax also tend to boost aggregate savings, and hence also contributes to the increase of public debt. Ultimately, the middle class suffers from the higher progressivity and does not benefit much from the lower capital tax, explaining that they have the lowest weights in the population.

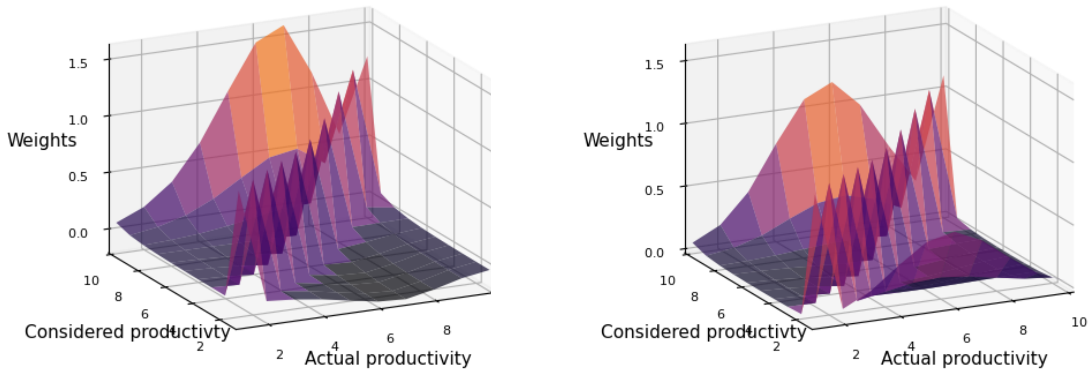
This increase in public debt and the decrease in capital taxes requires an increase in the labor tax to compensate for the loss in tax returns – as this instrument is adjusted to keep the public spending-to-GDP ratio unchanged.

6.5 Identification of IWFs

We now derive the estimated IWFs from the estimated SWF for each country. We follow the algorithm of Section 5.6, which relies on the estimation strategy of 3 of Section 4.4. The first step is to estimate the political weights ω_P of the various groups of agents. To approximate them, we rely on voter turnout in major elections as a function of average income. We thus follow the political economy literature, which uses the turnout inequality as a proxy for the evolution of political inequality (see Cage, 2024 for a discussion). Other proxies, such as donations in the US, are also considered in the political economy literature. However, political donations are very small in France, and do not allow for a proper comparison between France and the US.

Figure 6 plots the participation rate in percent as a function of the annual income in the US (panel a) and in France (panel b). For each country, the annual income is expressed as a percentage of the average annual income of the country.[28] In the US, average income is $\$\,51726$,

[28]US data are taken from Table 8 the Current Population Survey of November 2008 of the U.S. Census Bureau.



(a) US (b) France

Figure 7: IWF weights (measured as politically-weighted factors $(\omega_{P,y}\pi_y\omega_{y,\tilde{y}})_{y,\tilde{y}}$) for the US and France. *Actual productivity* is the productivity y of the agents, while *Considered productivity* is the productivity $\tilde{y}$ of the agents under consideration.

while it is 31093€ for France (both in 2007). We observe that the participation rate is increasing and concave in income for both countries. Our identification assumption is that the ratio of participation rates between income groups identifies the ratio of political weights. Formally: $\frac{\omega_{P,y}}{\omega_{P,\tilde{y}}} = \frac{Part_y}{Part_{\tilde{y}}}$, where $Part_y$ is the participation rate of income group y, which is interpolated from Figure 6. As a consequence, the shape of political weights follow the shape of participation rates. The higher average participation rate in France compared to the US is not reflected in political weights.

With these political weights, we can calculate the IWFs of equation (22) in Definition 3. The results are plotted in Figure 7, where the left panel (a) is for the US and the right panel (b) is for France. More precisely, we report the politically-weighted factors given by agents of productivity y (*Actual productivity*) to agents of productivity $\tilde{y}$ (*Considered productivity*). The politically-weighted factors on the z-axis are equal to the value of $\omega_{P,y}\pi_y\omega_{y,\tilde{y}}$ for all $y, \tilde{y} = 1, \ldots, 10$. As explained in Section 4.4, these factors have a simple interpretation: The larger the factor factors $\omega_{P,y}\pi_y\omega_{y,\tilde{y}}$, the more agents y affect the valuation of the planner for agents $\tilde{y}$.

First, in both countries, the diagonal features high weights, reflecting that the self-interest motive dominates the altruistic one: Agents mostly care about their own productivity.[29] The diagonal weights are also increasing with productivity in both countries, which mirrors the higher political weight of high-productivity agents. In the US, the increase is steeper than in France and diagonal weights reach higher values than in France, because SWF weights are also higher. In consequence, the most productive agents have the highest impact in social preferences, and this is especially true for the US. This is consistent with the results of Section 6.2.

Second, out of the diagonal, the US weights exhibit an increasing pattern in $\tilde{y}$ for each productivity level y. This is especially true for middle class agents, corresponding to intermediate

French data are taken from IPSOS data for participation rates as a function of occupations and DADS 2007 to obtain the annual income for each occupation.

[29] We recall that equation (22) implies $\pi_y\omega_{P,y}\hat{\omega}_{yy} = 1 + \frac{\pi_y(\omega_{P,y})^2}{\sum_{\tilde{y}\in\mathcal{Y}^\infty}\pi_{\tilde{y}}(\omega_{P,\tilde{y}})^2}(\omega_y - 1)$.

values of y, who have the largest share in the population. Such a shape has been qualified of Libertarian for the welfare functions in Section 3.4: higher weights are attributed to the most productive agents. Third, out of the diagonal, the French weights exhibit a U-shape pattern, consistent with the finding of 6.2 for the French SWF. Hence, the French welfare functions are Egalitarian for low level of productivity but Libertarian for high productivity levels.

7 Conclusion

We propose a methodology to identify the Social Welfare Function (SWF) and Individual Welfare Functions (IWFs) from the empirical wealth and income distributions and the actual tax structure. We implement it both for France and the US. Using four fiscal instruments – consumption, capital and progressive labor taxes, and public debt – we have estimated the SWFs in the two countries and showed that they differ from each other. The SWF for France puts a higher weight to low-productivity agents and is less heterogeneous than that of the US, while the US SWF has an increasing shape in productivity with larger weights given to higher-productivity agents. The US thus appear to be more Libertarian than France, while France is more Egalitarian than the US, especially for low income levels. These results pave the way for future research, particularly regarding the stability of social preferences over time. A key first step in this area is to investigate the role of the SWF in the fiscal response to economic shocks, particularly in terms of business cycle stabilization. Understanding this is essential for identifying the SWF by extracting insights from short-term changes in the fiscal system.

References

Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching up with the Joneses," *The American Economic Review*, 80, 38–42.

Açikgöz, O., M. Hagedorn, H. Holter, and Y. Wang (2022): "The Optimum Quantity of Capital and Debt," Working Paper, University of Oslo.

Acemoglu, D. (2010): *Political Economy*, Lecture Notes, MIT course.

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

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

Almås, I., A. W. Cappelen, and B. Tungodden (2020): "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy*, 128, 1753–1788.

Arrow, K. (1951): *Social Choice and Individual Values*, New York: John Wiley.

Bargain, O. and C. Keane (2010): "Tax-Benefit-Revealed Redistributive Preferences Over Time: Ireland 1987-2005," *Labour*, 24, 141–167.

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

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2021): "Inequality, Business Cycles, and Monetary-Fiscal Policy," *Econometrica*, 89, 2559–2599.

Billingsley, P. P. (2012): *Probability and Measure*, Wiley-Blackwell, Anniversary Edition.

Bourguignon, F. and S. Amadeo (2015): "Tax-Benefit Revealed Social Preferences," *Journal of Economic Inequality*, 1, 75–108.

Cage, J. (2024): "Political Inequality," *Annual Review of Economics,*, 16, 55–490.

Campbell, J. Y. and J. Cochrane (1999): "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

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

Chang, B., Y. Chang, and S.-B. Kim (2018): "Pareto Weights in Practice : Income Inequality and Tax Reform," *Review of Economic Dynamics*, 28, 181–204.

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

Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea after All!" *American Economic Review*, 99, 25–48.

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, 2431–2467.

Dyrda, S. and M. Pedroni (2022): "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks," *Review of Economic Studies*, 90, 744–780.

Fehr, E., D. Rutzler, and M. Sutter (2013): "The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence," *European Economic Review*, 64, 369–383.

Fehr, E. and K. Schmidt (2006): "The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories," Elsevier, vol. 1, chap. 08, 615–691, 1 ed.

Ferey, A., B. B. Lockwood, and D. Taubinsky (2024): "Sufficient Statistics for Nonlinear Tax Systems with General Across-Income Heterogeneity," *American Economic Review*, 114, 3206–49.

Ferriere, A. and G. Navarro (2023): "The Heterogeneous Effects of Government Spending: It's All About Taxes," *Review of Economic Studies*, forthcoming.

Finkelstein, A. and N. Hendren (2020): "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, 34, 146–167.

Fonseca, R., F. Langot, P.-C. Michaud, and T. Sopraseuth (2023): "Understanding Cross-Country Differences in Health Status and Expenditures: Health Prices Matter," *Journal of Political Economy*, 131, 1949–1993.

Gaertner, W. and E. Schokkaert (2012): *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*, Cambridge University Press.

Harsanyi, J. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, 309–321.

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

Heathcote, J. and H. Tsujiyama (2021): "Optimal Income Taxation: Mirrlees Meets Ramsey," *Journal of Political Economy*, 129, 3141–3184.

Hendren, N. (2020): "Measuring Economic Efficiency Using Inverse-Optimum Weights," *Journal of Public Economics*, 187.

Hendren, N. and B. Sprung-Keyser (2020): "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 3, 1209–1319.

Kaplow, L. and S. Shavell (2001): "Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle," *Journal of Political Economy*, 109, 281–286.

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

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

——— (2022b): "Refining the Truncation Method to Solve Heterogeneous-Agent Models," *Annals of Economics and Statistics*, 146, 65–91.

——— (2023): "Optimal Policies with Heterogeneous Agents: Truncation and Transitions," *Journal of Economic Dynamics and Control*, 156, 104737.

——— (2024): "Optimal Fiscal Policy with Heterogeneous Agents and Capital: Should We Increase or Decrease Public Debt and Capital Taxes?" *Journal of Political Economy*, forthcoming.

Lockwood, B. and M. Weinzierl (2016): "Positive and normative judgements implicit in US tax policy and the costs of unequal growth and recessions," *Journal of Monetary Economics*, 77, 30–47.

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

Piacquadio, P. G. (2017): "A Fairness Justification of Utilitarianism," *Econometrica*, 85, 1261–1276.

Rawls, J. (1971): *A Theory of Justice*, Harvard University Press.

Saez, E. and S. Stantcheva (2016): "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 106.

Sen, A. (1970): "The Impossibility of a Paretian Liberal," *Journal of Political Economy*, 78, 152–157.

——— (1977): "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," *Philosophy and Public Affairs*, 6, 317–344.

Sher, i. (2024): "Generalized Social Marginal Welfare Weights Imply Inconsistent Comparisons of Tax Policies," *American Economic Review*, Forthcoming.

Smith, A. (1759): *A Theory of Moral Sentiments*, Strand & Edinburgh: A. Millar; A. Kincaid & J. Bell.

Stantcheva, S. (2021): "Understanding Tax Policy: How do People Reason?" *The Quarterly Journal of Economics*, 136, 2309–2369.

Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics letters*, 20, 177–181.

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

Appendix

A Proofs related to the SWF construction

A.1 Construction of the measure on the set of idiosyncratic histories $\mathcal{Y}^\infty$

We can construct a probability space related to the set of infinite idiosyncratic histories, $\mathcal{Y}^\infty$. We summarize here the construction and further details can be found in LeGrand and Ragot (2022a, Appendix B.3). Consistently with the main text, we will typically denote by y^t an element of $\mathcal{Y}^\infty$. Such an element $\tilde{y}^t \in \mathcal{Y}^\infty$ can be described as a left-infinite sequence:

$$\tilde{y}^t = (\ldots, y_{-k}(\tilde{y}^t), \ldots, y_{-1}(\tilde{y}^t), y_0(\tilde{y}^t)),$$

where each $y_{-k} : \mathcal{Y}^\infty \to \mathcal{Y}$ is a coordinate function returning the idiosyncratic state k periods ago. For the sake of simplicity, and as in the main text, we will denote by $y_s^t := y_{-(t-s)}(y^t)$ for any $s \leq t$ the state in date s, which is $t - s$ periods ahead of date t.

We define $L(y^t)$ the past of history y^t – which discards the current state y_t^t:

$$L(y^t) = (\ldots, y_{t-k}^t, \ldots, y_{t-1}^t).$$

Consistently with the main text, we also denote by $y^{t-1,t}$ the past history of y^t: $y^{t-1,t} = L(y^t)$.

We can then define the cylinder sets $C_k(A)$ for any $k \geq 1$ and any $A \subset \mathcal{Y}^k$ as:

$$C_k(A) = \{y^t \in \mathcal{Y}^\infty : (y_{t-k+1}^t, \ldots, y_t^t) \in A\}.$$

The cylinder set $C_k(A)$ is the subset of $\mathcal{Y}^\infty$ containing all idiosyncratic histories whose truncation of length k belongs to A. We then define $\mathcal{C}_0$ as the set of all cylinder sets, which can be shown to be a field (Billingsley, 2012, Section 2). We denote by $\mathcal{F} := \sigma(\mathcal{C}_0)$ the cylindrical σ-algebra generated by $\mathcal{C}_0$, and we define the set function $\mu : \mathcal{C}_0 \to \mathbb{R}$ from the transition matrix Π and the stationary vector π, such that for any $k \geq 2$ and any $A \subset \mathcal{Y}^k$:

$$\begin{cases} \mu(C_k(A)) = \sum_{(y_{-k+1},\ldots,y_0)\in A} \pi_{y_{-k+1}} \Pi_{y_{-k+1}y_{-k+2}} \ldots \Pi_{y_{-1}y_0} & \text{for any } k \geq 2 \text{ and } A \subset \mathcal{Y}^k, \\ \mu(C_1(A)) = \sum_{y_0 \in A} \pi_{y_0} & \text{for any } A \subset \mathcal{Y}. \end{cases} \tag{48}$$

Finally, we can state the following lemma.

Lemma 1 *The triplet* $(\mathcal{Y}^\infty, \mathcal{F}, \mu)$ *is a probability space.*

Proof. A proof can be found in Billingsley (2012, Section 2). The key part of the proof is to extend the measure μ defined on $\mathcal{C}_0$ to a measure defined on $\sigma(\mathcal{C}_0) = \mathcal{F}$. ■ A consequence of Lemma 1 is that $\mu(\mathcal{Y}^\infty) = 1$, or $\int_{y^t \in \mathcal{Y}^\infty} \mu(dy^t) = 1$.

A.2 The conditional measure

To lighten formulas, for any $(y_{-k+1}, \ldots y_0) \in \mathcal{Y}^k$, we define $y_{-k+1:0} := (y_{-k+1}, \ldots, y_0)$ the vector of length k, containing elements whose indices range from $-k+1$ to 0. Similarly, for any $\tilde{y}^t \in \mathcal{Y}^\infty$,

$\tilde{y}^t_{(t-k+1:t)} = (\tilde{y}^t_{t-k+1}, \dots, \tilde{y}^t_t)$ is the vector of the k last realization of $\tilde{y}^t$. We define μ_1 by:

$$\mu_1(C_1(Y_1)|\tilde{y}^t) = \sum_{y_{t+1}\in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}}, \text{ for any } Y_1 \subset \mathcal{Y}, \tag{49}$$

$$\forall k \geq 2,\ \mu_1(C_k(Y_k)|\tilde{y}^t) = \sum_{y_{(t-k+2):t+1}\in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t}=\tilde{y}^t_{t-k+2:t}}, \text{ for any } Y_k \subset \mathcal{Y}^k, \tag{50}$$

where for any elements $x, \tilde{x}$ of the same set, $1_{x=\tilde{x}} = 1$ if $x = \tilde{x}$ and $1_{x=\tilde{x}} = 0$ otherwise. Intuitively, the expression in (50) sums over all possible vectors $y_{(t-k+2):t+1}$ of length k, the probability to switch from $\tilde{y}^t$ to an history ending up in $y_{(t-k+2):t+1}$. The latter probability is equal to the probability to switch from $\tilde{y}^t_t$ to y_{t+1}, provided that $\tilde{y}^t$ and $y_{(t-k+2):t}$ are compatible (i.e., the $k-1$ last realization of $\tilde{y}^t$ equals $y_{(t-k+2):t}$). Note that we could also write: $1_{y_{(t-k+2:t)}=\tilde{y}^t_{(t-k+2:t)}} = \prod_{j=0}^{k-2} 1_{y_{t-j}=\tilde{y}^t_{t-j}}$.

Lemma 2 *For any $\tilde{y}^t \in \mathcal{Y}^\infty$, the set function $C \in \mathcal{C}_0 \mapsto \mu_1(C|\tilde{y}^t)$ is a pre-measure.*

Proof.

In the remainder of the proof, we set $\tilde{y}^t \in \mathcal{Y}^\infty$. As a preliminary, we state two properties that we will use extensively below:

- For all $k' \geq k \geq 1$, for all $Y_k \subset \mathcal{Y}^k$, for all $Y'_{k'-k} \subset \mathcal{Y}^{k'-k}$:

$$\sum_{y_{(t-k'+2):t+1}\in Y'_{k'-k}\times Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t}=\tilde{y}^t_{(t-k'+2):t}} = \sum_{y_{(t-k+2):t+1}\in Y_k} \Bigg(\Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}} \times \sum_{y_{(t-k'+2):(t-k+1)}\in Y'_{k'-k}} 1_{y_{(t-k'+2):(t-k+1)}=\tilde{y}^t_{(t-k'+2):(t-k+1)}}\Bigg) \tag{51}$$

- For all $k' \geq k \geq 0$:

$$\sum_{y_{t-k':t-k}\in\mathcal{Y}^{k'-k+1}} 1_{y_{t-k':t-k}=\tilde{y}^t_{t-k':t-k}} = 1. \tag{52}$$

In the remainder they will be referred to by their equation numbering. The proof of (52) is straightforward and comes from the fact that $(\tilde{y}^t_{t-k'}, \dots, \tilde{y}^t_{t-k})$ is a unique element of the set $\mathcal{Y}^{k'-k+1}$. For the proof of (51), we denote by $S_{k,k'}$ the left hand side. We have:

$$S_{k,k'} = \sum_{y_{(t-k+2):t+1}\in Y_k} \sum_{y_{(t-k'+2):(t-k+1)}\in Y'_{k'-k}} \Bigg(\Pi_{\tilde{y}^t_t y_{t+1}} \times 1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}} 1_{y_{(t-k'+2):(t-k+1)}=\tilde{y}^t_{(t-k'+2):(t-k+1)}}\Bigg) \tag{53}$$

where we have used the properties of a sum on a product space and the fact that $(x, y) = (x', y')$ iff $x = x'$ and $y = y'$ (where (x, x') and (y, y') are two pairs of vectors of the same length). Then we can factorize $\Pi_{\tilde{y}^t_t y_{t+1}}$ and $1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}}$ in (53), as they are independent from the sum over $y_{(t-k'+2):(t-k+1)} \in Y'_{k'-k}$. We then readily obtain (51).

We now go back to the proof of Lemma 2. Three properties need to hold: (i) well-defined, (ii) (countably) additive, (iii) $\mu_1(\mathcal{Y}^\infty|\tilde{y}^t) = 1$ for all $\tilde{y}^t \in \mathcal{Y}^\infty$.

For Point (i), we need to check that $\mu_1(C_k(Y_k)|\tilde{y}^t) = \mu_1(C_{k'}(\mathcal{Y}^{k'-k} \times Y_k)|\tilde{y}^t)$ for all $k' \geq k$. We have for $k \geq 2$:

$$\begin{aligned}\mu_1(C_{k'}(\mathcal{Y}^{k'-k} \times Y_k)|\tilde{y}^t) &= \sum_{y_{(t-k'+2):t+1} \in \mathcal{Y}^{k'-k} \times Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}}, \\ &= \sum_{y_{(t-k+2):t+1} \in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}} = \mu_1(C_k(Y_k)|\tilde{y}^t), \end{aligned} \tag{54}$$

where the first and last equalities use the definition of μ_1, and the second the combination of properties (51) and (52).

We need to prove the result for $k = 1$. We have for $k' \geq 1$:

$$\begin{aligned}\mu_1(C_{k'}(\mathcal{Y}^{k'-1} \times Y_1)|\tilde{y}^t) &= \sum_{y_{t+1} \in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} \Bigg(\sum_{y_{(t-k'+2):t} \in \mathcal{Y}^{k'-1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}} \Bigg), \\ &= \sum_{y_{t+1} \in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} = \mu_1(C_1(Y_1)|\tilde{y}^t),\end{aligned}$$

where the first equality combines the definition of μ_1 and (51), the second uses the property (52), and the last the definition of μ_1.

For point (ii), we consider two disjoint cylinders, $C_k(Y_k)$ and $C_{k'}(Y'_{k'})$ ($k' \geq k$, $Y_k \subset \mathcal{Y}^k$ and $Y'_{k'} \subset \mathcal{Y}^{k'}$). Since both cylinders are disjoint, then $\mathcal{Y}^{k'-k} \times Y_k$ and $Y'_{k'}$ are disjoint too. We deduce that:

$$\begin{aligned}\mu_1(C_k(Y_k) \cup C_{k'}(Y'_{k'})|\tilde{y}^t) &= \mu_1(C_{k'}((\mathcal{Y}^{k'-k} \times Y_k) \cup Y'_{k'})|\tilde{y}^t), \\ &= \sum_{y_{(t-k'+2):t+1} \in (\mathcal{Y}^{k'-k} \times Y_k) \cup Y'_{k'}} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}}, \\ &= \sum_{y_{(t-k'+2):t+1} \in (\mathcal{Y}^{k'-k} \times Y_k)} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}} \\ &+ \sum_{y_{(t-k'+2):t+1} \in Y'_{k'}} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}}, \\ &= \sum_{y_{(t-k+2):t+1} \in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}} \\ &+ \sum_{y_{(t-k'+2):t+1} \in Y'_{k'}} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k'+2):t} = \tilde{y}^t_{(t-k'+2):t}}, \\ &= \mu_1(C_k(Y_k)|\tilde{y}^t) + \mu_1(C_{k'}(Y'_{k'})|\tilde{y}^t),\end{aligned}$$

where the first equality uses the algebra property of cylinder sets, the second the definition of μ_1, the third the property that $\mathcal{Y}^{k'-k} \times Y_k$ and $Y'_{k'}$ are disjoint, the fourth the combination of properties (51) and (52), and the last the definition of μ_1 twice. We have thus proved that $\mu_1(\cdot|\tilde{y}^t)$ is finitely additive.

For point (iii), let $k \geq 2$. We have:

$$\mu_1(C_k(\mathcal{Y}^k)|\tilde{y}^t) = \sum_{y_{t+1}\in\mathcal{Y}} \Pi_{\tilde{y}^t_t y_{t+1}} \sum_{y_{(t-k+2):t}\in\mathcal{Y}^{k-1}} 1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}},$$
$$= \sum_{y_{t+1}\in\mathcal{Y}} \Pi_{\tilde{y}^t_t y_{t+1}} = 1,$$

where the first equality uses the definition of μ_1 and (51), the second property (52), and the third the property of the transition matrix Π. We thus deduce that $\mu_1(\mathcal{Y}^\infty|\tilde{y}^t) = 1$.

We have proven that $\mu_1(\cdot|\tilde{y}^t)$ is a finitely additive probability measure on the algebra $\mathcal{C}_0$. Billingsley (2012, Theorem 2.3) states that any finitely additive probability measure on the cylinder algebra is countably additive. We thus conclude that μ_1 is a countably additive probability measure on $\mathcal{C}_0$ and is thus a pre-measure on $\mathcal{C}_0$. ■

We then prove the following lemma. We recall that $\mathcal{F}$ the cylindrical σ-algebra, $\sigma(\mathcal{C}_0)$, generated by $\mathcal{C}_0$.

Lemma 3 *For all $\tilde{y}^t \in \mathcal{Y}^\infty$, the function $\mu_1(\cdot|\tilde{y}^t)$ uniquely extends to a measure on $\mathcal{F}$.*

The proof is similar to the one showing the extension of μ as a pre-measure on $\mathcal{C}_0$ to a measure on $\mathcal{F}$. It relies on the Hahn-Kolmogorov theorem (Billingsley 2012, Theorem 3.1). See LeGrand and Ragot (2022a, Lemma 3 in Section B.3).

Finally we can state the following lemma, showing that μ_1 is a conditional measure.

Lemma 4 *For all $\tilde{y}^t \in \mathcal{Y}^\infty$ and for all $F \in \mathcal{F}$:*

$$\int_{\tilde{y}^t\in\mathcal{Y}^\infty} \mu_1(F|\tilde{y}^t)\mu(d\tilde{y}^t) = \mu(F) \tag{55}$$

.

Proof. We first prove (55) for F being a cylinder set.

First, let $Y_1 \subset \mathcal{Y}$ and consider $C_1(Y_1)$. We have:

$$\int_{\tilde{y}^t\in\mathcal{Y}^\infty} \mu_1(C_1(Y_1)|\tilde{y}^t)\mu(d\tilde{y}^t) = \int_{\tilde{y}^t\in\mathcal{Y}^\infty} \sum_{y_{t+1}\in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} \mu(d\tilde{y}^t),$$
$$= \sum_{\tilde{y}^t_t\in\mathcal{Y}} \pi_{\tilde{y}^t_t} \sum_{y_{t+1}\in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}},$$
$$= \sum_{y_{t+1}\in Y_1} \sum_{\tilde{y}^t_t\in\mathcal{Y}} \pi_{\tilde{y}^t_t} \Pi_{\tilde{y}^t_t y_{t+1}},$$
$$= \sum_{y_{t+1}\in Y_1} \pi_{y_0} = \mu(C_1(Y_1)),$$

where the first equality comes from the definition (49) of μ_1, the second from the fact that the integral is actually carried over a cylinder set of the form $C_1(\mathcal{Y})$ and from the definition (48) of μ, the third from the permutation of finite sums, the fourth from the fact that π is stationary ($\sum_{y\in\mathcal{Y}} \pi_y \Pi_{yy'} = \pi_{y'}$), and the last from the definition (48) of μ.

Second, let $Y_k \subset \mathcal{Y}^k$ and consider $C_k(Y_k)$. We have:

$$
\begin{aligned}
&\int_{\tilde{y}^t \in \mathcal{Y}^\infty} \mu_1(C_k(Y_k)|\tilde{y}^t)\mu(d\tilde{y}^t) \\
&= \int_{\tilde{y}^t \in \mathcal{Y}^\infty} \sum_{y_{(t-k+2):t+1} \in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}} \mu(d\tilde{y}^t), \\
&= \sum_{\tilde{y}^t_{(t-k+2):t} \in \mathcal{Y}^{k-1}} \pi_{\tilde{y}^t_{t-k+2}} \Pi_{\tilde{y}^t_{t-k+2}\tilde{y}^t_{t-k+3}} \dots \Pi_{\tilde{y}^t_{t-1}\tilde{y}^t_t} \sum_{y_{(t-k+2):t+1} \in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}}, \\
&= \sum_{y_{(t-k+2):t+1} \in Y_k} \sum_{\tilde{y}^t_{(t-k+2):t} \in \mathcal{Y}^{k-1}} \pi_{\tilde{y}^t_{t-k+2}} \Pi_{\tilde{y}^t_{t-k+2}\tilde{y}^t_{t-k+3}} \dots \Pi_{\tilde{y}^t_{t-1}\tilde{y}^t_t} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}}, \\
&= \sum_{y_{(t-k+2):t+1} \in Y_k} \pi_{y_{t-k+2}} \Pi_{y_{t-k+2}y_{t-k+3}} \dots \Pi_{y_{t-1}y_t} \Pi_{y_t y_{t+1}} = \mu(C_k(Y_k)),
\end{aligned}
$$

where the first equality comes from the definition (50) of μ_1, the second from the fact that the integral is actually carried over a cylinder set of the form $C_{k-1}(\mathcal{Y}^{k-1})$ and from the definition (48) of μ, the third from the permutation of finite sums, the fourth from the fact that all terms in the sum over $\tilde{y}^t_{(t-k+2):t} \in \mathcal{Y}^{k-1}$ are zero but the one for $\tilde{y}^t_{(t-k+2):t} = y_{(t-k+2):t}$, and the last from the definition (48) of μ. This proves (55) on $\mathcal{C}_0$. Since cylinder sets are a generating family of $\mathcal{F}$ and a π-system, the equality (55) also holds on $\sigma(\mathcal{C}_0) = \mathcal{F}$. ■

We finally prove the following lemma. It justifies that in the main text, we define μ_1 as $\mu_1(dy^{t+1}|\tilde{y}^t) = \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(dy^t)$.

Lemma 5 *For all $\tilde{y}^t \in \mathcal{Y}^\infty$ and for all $C \in \mathcal{C}_0$, we have:*

$$\mu_1(C|\tilde{y}^t) = \int_{y^{t+1} \in C} \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(L(dy^{t+1})),$$

where $\delta_{\tilde{y}^t}$ is the Dirac mass in $\tilde{y}^t$.

Proof. Let $Y_k \subset \mathcal{Y}^k$ for some $k \geq 1$.

For $k = 1$, we have:

$$
\begin{aligned}
\int_{y^{t+1} \in C_1(Y_1)} \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(L(dy^{t+1})) &= \int_{(y^t, y_{t+1}) \in \mathcal{Y}^\infty \times Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} \delta_{\tilde{y}^t}(dy^t), \\
&= \sum_{y_{t+1} \in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} \int_{y^t \in \mathcal{Y}^\infty} \delta_{\tilde{y}^t}(dy^t), \\
&= \sum_{y_{t+1} \in Y_1} \Pi_{\tilde{y}^t_t y_{t+1}} = \mu_1(C_1(Y_1)|\tilde{y}^t),
\end{aligned}
$$

where the first equality comes from using $C_1(Y_1) = \mathcal{Y}^\infty \times Y_1$, the second from using the Fubini theorem (we consider σ-finite measure spaces and $(\tilde{y}^t, y_{t+1}) \in \mathcal{Y}^\infty \times Y_1 \mapsto \Pi_{\tilde{y}^t_t y_{t+1}} \delta_{\tilde{y}^t}(dy^t)$ is integrable), the third the property of Dirac mass, and fourth the definition of μ_1.

For $k \geq 2$, we have:

$$\int_{y^{t+1} \in C_k(Y_k)} \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(L(dy^{t+1})) \tag{56}$$

$$= \int_{(y^{t-k+1}, y_{(t-k+2):t+1}) \in \mathcal{Y}^\infty \times Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t} = \tilde{y}^t_{(t-k+2):t}} \delta_{L^{(k-1)}(\tilde{y}^t)}(dy^{t-k+1}), \tag{57}$$

where $L^{(k-1)}(\cdot)$ is k-iterate of the shift operator L: $L^{(k-1)}(\tilde{y}^t) = (\ldots, \tilde{y}^t_{t-k}, \tilde{y}^t_{t-k+1})$. To write equality (56), we have used that $C_k(Y_k) = \mathcal{Y}^\infty \times Y_k$ and the fact that $\delta_{\tilde{y}^t}(Y_{k-1} \times C) = \sum_{y_{(t-k+2):t} \in Y_{k-1}} 1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}} \delta_{L^{(k-1)}(\tilde{y}^t)}(C)$ for all $C \in \mathcal{C}_0$ and $Y_{k-1} \subset \mathcal{Y}^{k-1}$. Hence the two measures coincide on $\mathcal{C}_0$ and are also σ- finite. Using Fubini and the property of a Dirac mass, we obtain from (56):

$$\int_{y^{t+1} \in C_k(Y_k)} \Pi_{\tilde{y}^t_t y^{t+1}_{t+1}} \delta_{\tilde{y}^t}(L(dy^{t+1})) = \sum_{y_{(t-k+2):t+1} \in Y_k} \Pi_{\tilde{y}^t_t y_{t+1}} 1_{y_{(t-k+2):t}=\tilde{y}^t_{(t-k+2):t}},$$
$$= \mu_1(C_k(Y_k)|\tilde{y}^t),$$

where the last equality comes from the definition of μ_1. This concludes the proof. ■

A.3 Proof of Equality 14

We consider an agent whose individual preferences are represented by a utility function $V : \mathcal{Y}^\infty \to \mathbb{R}$. The allocation is subsumed. Let $y^t \in \mathcal{Y}^\infty$. The representation result of equation (13) becomes, after splitting the sum for $s = 0$ and $s \geq 1$:

$$V(y^t) = U(y^t) + \beta \sum_{s=1}^{\infty} \int_{y^{t+s} \in \mathcal{Y}^\infty} \beta^{s-1} U(y^{t+s}) \mu_s(dy^{t+s}|y^t). \tag{58}$$

Using the definition of μ_s given in Section 4.1 and Bayes rule, we have, for all $s \geq 1$:

$$\mu_s(dy^{t+s}|y^t) = \int_{y^{t+1} \in \mathcal{Y}^\infty} \mu_{s-1}(dy^{t+s}|y^{t+1}) \mu_1(dy^{t+1}|y^t). \tag{59}$$

In words, it means that the probability of transitioning from y^t to y^{t+s} in s periods is equal to the product of the probabilities of transitioning from y^t to y^{t+1} (in 1 period) and from y^{t+1} to y^{t+s} (in $s-1$ periods), summed over all possible histories y^{t+1}. Using (59), the expression of V in (58) becomes:

$$V(y^t) = U(y^t) + \beta \sum_{s=1}^{\infty} \int_{y^{t+s} \in \mathcal{Y}^\infty} \beta^{s-1} U(y^{t+s})$$
$$\times \int_{y^{t+1} \in \mathcal{Y}^\infty} \mu_{s-1}(dy^{t+s}|y^{t+1}) \mu_1(dy^{t+1}|y^t).$$

Since all function are integrable, we can use the Fubini theorem twice, to swap the order of integrals and then of the sum and of the integral on $y^{t+1} \in \mathcal{Y}^\infty$. We obtain:

$$V(y^t) = U(y^t) + \tag{60}$$
$$+\beta \int_{y^{t+1} \in \mathcal{Y}^\infty} \left\{ \sum_{s=1}^{\infty} \int_{y^{t+s} \in \mathcal{Y}^\infty} \beta^{s-1} U(y^{t+s}) \mu_{s-1}(dy^{t+s}|y^{t+1}) \right\} \mu_1(dy^{t+1}|y^t).$$

The term between curly braces is $V(y^{t+1})$ using (13) applied to $t+1$. The previous equality becomes:

$$V(y^t) = U(y^t) + \beta \int_{y^{t+1} \in \mathcal{Y}^\infty} V(y^{t+1}) \mu_1(dy^{t+1}|y^t).$$

which gives the representation (14) using the notation $\mathbb{E}_{y^{t+1}}\left[V(y^{t+1})|y^t\right] = \int_{y^{t+1}\in\mathcal{Y}^\infty} V(y^{t+1})\mu_1(dy^{t+1}|y^t)$. for the conditional expectation.

A.4 Proof of Proposition 1

Using the definition (16), the expression (17) of the SWF becomes:

$$SWF = \int_{y^t\in\mathcal{Y}^\infty} \omega_P(y^t) \int_{\tilde{y}^t\in\mathcal{Y}^\infty} \hat{V}(y^t, \tilde{y}^t)\mu(d\tilde{y}^t)\mu(dy^t),$$

which can be further simplified using the expression (15) of $\hat{V}$:

$$\begin{aligned} SWF &= \int_{y^t\in\mathcal{Y}^\infty} \omega_P(y^t) \int_{\tilde{y}^t\in\mathcal{Y}^\infty} \sum_{s=0}^{\infty} \int_{\hat{y}^{t+s}\in\mathcal{Y}^\infty} \beta^s \hat{\omega}(y^t, \hat{y}^{t+s}) U(\hat{y}^{t+s}) \mu_s(d\hat{y}^{t+s}|\tilde{y}^t)\mu(d\tilde{y}^t)\mu(dy^t). \\ &= \int_{y^t\in\mathcal{Y}^\infty} \omega_P(y^t) \sum_{s=0}^{\infty} \int_{\tilde{y}^t\in\mathcal{Y}^\infty} \int_{\hat{y}^{t+s}\in\mathcal{Y}^\infty} \beta^s \hat{\omega}(y^t, \hat{y}^{t+s}) U(\hat{y}^{t+s}) \mu_s(d\hat{y}^{t+s}|\tilde{y}^t)\mu(d\tilde{y}^t)\mu(dy^t). \end{aligned}$$

Since all functions under consideration are positive and measurable and since it is assumed that $SWF < \infty$, we can use the Fubini theorem to permute the order of integrals and obtain:

$$\begin{aligned} SWF = \sum_{s=0}^{\infty} \int_{\hat{y}^{t+s}\in\mathcal{Y}^\infty} \beta^s &\left[\int_{y^t\in\mathcal{Y}^\infty} \omega_P(y^t)\hat{\omega}(y^t, \hat{y}^{t+s})\mu(dy^t)\right] U(\hat{y}^{t+s}) \\ &\times \left[\int_{\tilde{y}^t\in\mathcal{Y}^\infty} \mu_s(d\hat{y}^{t+s}|\tilde{y}^t)\mu(d\tilde{y}^t)\right]. \end{aligned}$$

A straightforward extension of Lemma 4 for μ_s (for any $s \geq 1$) yields: $\int_{\tilde{y}^t\in\mathcal{Y}^\infty} \mu_s(d\hat{y}^{t+s}|\tilde{y}^t)\mu(d\tilde{y}^t) = \mu(d\hat{y}^{t+s})$. Using the definition (19) of the weights ω, we deduce:

$$SWF = \sum_{s=0}^{\infty} \int_{\hat{y}^{t+s}\in\mathcal{Y}^\infty} \beta^s \omega(\hat{y}^{t+s}) U(\hat{y}^{t+s}) \mu(d\hat{y}^{t+s}),$$

which proves Proposition 1.

As a final remark, observe that by splitting the sum over s into $s = 0$ and a sum for $s \geq 1$, we also obtain the following expression for SWF: $SWF = \int_{y^t\in\mathcal{Y}^\infty} \omega(y^t)U(y^t, A)\mu(dy^t) + \beta\, SWF$.

A.5 Proof of Proposition 2

We recall the expression of the SWF when the allocation is explicit:

$$SWF(A) = \sum_{t=0}^{\infty} \beta^t \int_{y^t\in\mathcal{Y}^\infty} \omega(y^t)U(y^t, A)\mu(dy^t).$$

Let assume that the weights are non-negative and consider two allocations A and A' such that A element-wise dominates A'. We thus have $U(y^t, A) \geq U(y^t, A')$ for all y^t. The non-negativity of weights implies that: $\beta^t\omega(y^t)U(y^t, A) \geq \beta^t\omega(y^t)U(y^t, A')$ for all y^t, which after integration and sum yields $SWF(A) \geq SWF(A')$.

Let us assume that $SWF(A) \geq SWF(A')$ for any pair of allocations A and A' such that A element-wise dominates A'. Let us assume that the weights are strictly negative on a subset

$\mathcal{X} \subset \mathcal{Y}^\infty$ of positive measure. We consider an allocation A'. We construct the allocation A such that A and A' coincide on $\mathcal{Y}^\infty \setminus \mathcal{X}$ and A strictly dominates A' on $\mathcal{X}$. We thus have $U(y^t, A) = U(y^t, A')$ for all $y^t \in \mathcal{Y}^\infty \setminus \mathcal{X}$ and $U(y^t, A) > U(y^t, A')$ for all $y^t \in \mathcal{X}$: A element-wise dominates A'. We thus deduce that:

$$\begin{aligned}\int_{y^t\in\mathcal{Y}^\infty} \omega(y^t)(U(y^t,A)-U(y^t,A'))\mu(dy^t) &= \int_{y^t\in\mathcal{Y}^\infty\setminus\mathcal{X}} \omega(y^t)(U(y^t,A)-U(y^t,A'))\mu(y^t)\\ &+ \int_{y^t\in\mathcal{X}} \omega(y^t)(U(y^t,A)-U(y^t,A'))\mu(y^t),\\ &= \int_{y^t\in\mathcal{X}} \omega(y^t)(U(y^t,A)-U(y^t,A'))\mu(y^t),\\ &< 0,\end{aligned}$$

where the first equality is a split of the integral over two disjoint sets, the second comes from $U(y^t, A) = U(y^t, A')$ on $\mathcal{Y}^\infty \setminus \mathcal{X}$, and the third from $U(y^t, A) > U(y^t, A')$ and $\omega(y^t) < 0$ on $\mathcal{X}$.

Summing the previous discounted inequality implies $SWF(A) < SWF(A')$, which is a contradiction. We must thus have positive weights. This concludes the proof.

A.6 Proof of equation (22) in Definition 3

The program in Definition 3 is:

$$\begin{aligned}(\tilde{\omega}_{\tilde{y}y})_{\tilde{y},y} =& \text{argmin}_{(\hat{\omega}_{\tilde{y}y})_{\tilde{y},y}} \sum_{(y,\tilde{y})\in\mathcal{Y}^{\infty 2}} \pi_{\tilde{y}} \left(\hat{\omega}_{\tilde{y}y} - \frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y}\right)^2,\\ &\text{s.t. } \omega_y = \sum_{\tilde{y}\in\mathcal{Y}^\infty} \pi_{\tilde{y}}\omega_{P,\tilde{y}}\hat{\omega}_{\tilde{y}y} \ (y \in \mathcal{Y}).\end{aligned} \tag{61}$$

We denote by $2\lambda_y$ the Lagrange multiplier to the constraint of equation (61) for $y \in \mathcal{Y}$. We obtain the following Lagrangian:

$$\mathcal{L} = \frac{1}{2} \sum_{(y,\tilde{y})\in\mathcal{Y}^{\infty 2}} \pi_{\tilde{y}} \left(\hat{\omega}_{\tilde{y}y} - \frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y}\right)^2 - \sum_{y\in\mathcal{Y}} \lambda_y \left(\sum_{\tilde{y}\in\mathcal{Y}^\infty} \pi_{\tilde{y}}\omega_{P,\tilde{y}}\hat{\omega}_{\tilde{y}y} - \omega_y\right).$$

Computing the derivative with respect to $\hat{\omega}_{\tilde{y}y}$ yields the following FOC:

$$\hat{\omega}_{\tilde{y}y} = \frac{1_{y=\tilde{y}}}{\omega_{P,y}\pi_y} + \lambda_y\omega_{P,\tilde{y}}.$$

Using the constraint of equation (61), we deduce:

$$\lambda_y = \frac{\omega_{P,\tilde{y}}}{\sum_{\tilde{y}\in\mathcal{Y}^\infty} \pi_{\tilde{y}}(\omega_{P,\tilde{y}})^2}(\omega_y - 1),$$

which finally implies equation (A.6).

B Competitive equilibrium

We provide a formal definition of a competitive equilibrium.

Definition 4 (Competitive equilibrium) *A sequential competitive equilibrium is a collection of individual allocations* $(c_{i,t}, l_{i,t}, a_{i,t}, \nu_{i,t})_{t\geq 0, i\in\mathcal{I}}$*, of aggregate quantities* $(K_t, L_t, Y_t)_{t\geq 0}$*, of price processes* $(w_t, r_t, \tilde{w}_t, \tilde{r}_t)_{t\geq 0}$*, and of fiscal policies* $(\tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t)_{t\geq 0}$*, such that, for initial conditions and initial values of capital stock and public debt verifying* $K_{-1} + B_{-1} = \int_i a_{i,-1}\ell(di)$*, we have:*

1. *given prices, the functions* $(c_{i,t}, l_{i,t}, a_{i,t}, \nu_{i,t})_{t\geq 0, i\in\mathcal{I}}$ *solve the agent's optimization program in equations (30)–(32);*

2. *financial, labor, and goods markets clear at all dates: for any* $t \geq 0$*, equation (35) holds;*

3. *the government budget is balanced at all dates: equation (28) holds for all* $t \geq 0$*;*

4. *factor prices* $(w_t, r_t, \tilde{w}_t, \tilde{r}_t)_{t\geq 0}$ *are consistent with condition (24) and post-tax definitions (27).*

A steady-state competitive equilibrium is a competitive equilibrium for which the joint distribution of agents' decisions (c, l, a, ν), aggregate quantities K, L, Y, prices $w, r, \tilde{w}, \tilde{r}$, and fiscal policy $(\tau^c, \tau^K, \tau, \kappa, B)$ are time-invariant.

C The Ramsey program 3

C.1 Reformulating the Ramsey program

We now reformulate the Ramsey problem. We define the following variables:

$$\tilde{a}_{i,t} := \frac{a_{i,t}}{1+\tau_t^c}, \tag{62}$$

$$W_t := \frac{w_t}{1+\tau_t^c}, \tag{63}$$

$$R_t := \frac{(1+r_t)(1+\tau_{t-1}^c)}{1+\tau_t^c}, \tag{64}$$

which represents the asset choices in (62), the wage rate in (63), and the interest rate in (64). With this notation, the agent's budget and credit constraints become:

$$c_{i,t} + \tilde{a}_{i,t} = W_t(y_{i,t}l_{i,t})^{1-\tau_t} + R_t\tilde{a}_{i,t-1}, \tag{65}$$

$$\tilde{a}_{i,t} \geq -\frac{\overline{a}}{1+\tau_t^c} := -\tilde{\overline{a}}. \tag{66}$$

Since taxes and prices are considered as given by agents, we can equivalently state their optimization program using the notation (62)–(64) and the constraints (65) and (66), rather than the original notation and the constraints (31) and (32). This modifies Euler equations (33)–(34) as follows:

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

$$v'(l_{i,t}) = (1-\tau_t)W_t y_{i,t}(y_{i,t}l_{i,t})^{-\tau_t}u'(c_{i,t}).$$

We now turn to the governmental budget constraint. We further define:

$$\tilde{B}_t := \frac{B_t}{(1+\tau_t^c)}, \tag{67}$$

$$\tilde{A}_t := \frac{A_t}{1+\tau_t^c}. \tag{68}$$

and

$$\hat{B}_t := (1+\tau_t^c)\tilde{B}_t - \tau_t^c \tilde{A}_t, \tag{69}$$

With these new definitions, the financial market equilibrium given by (43) holds, as we have $\tilde{A}_t = \int_i \tilde{a}(i)l(di)$.

Using the government budget constraint defined in (28), we have:

$$G_t + (1+r_t)B_{t-1} + w_t \int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + r_t K_{t-1} = \tau_t^c C_t + F(K_{t-1}, L_t, s_t) + B_t.$$

Using the resource constraint $C_t + G_t + K_t = F(K_{t-1}, L_t, s_t) + K_{t-1}$, we obtain:

$$\begin{aligned}
&G_t + (1+r_t)B_{t-1} + w_t \int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + r_t K_{t-1} = \\
&\tau_t^c(F(K_{t-1}, L_t, s_t) - G_t - (K_t - K_{t-1})) + F(K_{t-1}, L_t, s_t) + B_t.
\end{aligned}$$

Divide both sides of the equation above by $(1+\tau_t^c)$ and obtain:

$$\begin{aligned}
&G_t + \frac{1+r_t}{1+\tau_t^c}B_{t-1} + \frac{w_t}{1+\tau_t^c}\int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + \frac{r_t}{1+\tau_t^c}K_{t-1} = \\
&-\frac{\tau_t^c}{1+\tau_t^c}(K_t - K_{t-1}) + F(K_{t-1}, L_t, s_t) + \frac{B_t}{1+\tau_t^c}.
\end{aligned}$$

Using the definitions (63), (64), and (67):

$$G_t + R_t\tilde{B}_{t-1} + W_t \int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + \frac{r_t}{1+\tau_t^c}K_{t-1} = -\frac{\tau_t^c}{1+\tau_t^c}(K_t - K_{t-1}) + F(K_{t-1}, L_t, s_t) + \tilde{B}_t.$$

We now substitute the expression of K_t and K_{t-1}. From (67) and (68), we have $K_{t-1} = A_{t-1} - B_{t-1} = (1+\tau_{t-1}^c)(\tilde{A}_{t-1} - \tilde{B}_{t-1})$ and:

$$\begin{aligned}
&G_t + R_t\tilde{B}_{t-1} + W_t \int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + \frac{r_t(1+\tau_{t-1}^c)}{1+\tau_t^c}(\tilde{A}_{t-1} - \tilde{B}_{t-1}) = \\
&\frac{\tau_t^c(1+\tau_{t-1}^c)}{1+\tau_t^c}(\tilde{A}_{t-1} - \tilde{B}_{t-1}) + F(K_{t-1}, L_t, s_t) - \tau_t^c(\tilde{A}_t - \tilde{B}_t) + \tilde{B}_t.
\end{aligned}$$

Observe from (64) and (69) that $\frac{r_t(1+\tau_{t-1}^c)}{1+\tau_t^c} = R_t - \frac{1+\tau_{t-1}^c}{1+\tau_t^c}$ and $-\tau_t^c(\tilde{A}_t - \tilde{B}_t) + \tilde{B}_t = \hat{B}_t$. This yields:

$$\begin{aligned}
&G_t + R_t\tilde{B}_{t-1} + W_t \int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di) + (R_t - (1+\tau_{t-1}^c))(\tilde{A}_{t-1} - \tilde{B}_{t-1}) = \\
&F(K_{t-1}, L_t, s_t) + \hat{B}_t.
\end{aligned}$$

Finally, using (69) in period $t-1$ (i.e., $\hat{B}_{t-1} = (1+\tau_{t-1}^c)\tilde{B}_{t-1} - \tau_{t-1}^c \tilde{A}_{t-1}$) we get:

$$G_t + W_t \int_i (y_{i,t} l_{i,t})^{1-\tau_t} \ell(di) + (R_t - 1)\tilde{A}_{t-1} + \hat{B}_{t-1} =$$
$$F(K_{t-1}, L_t, s_t) + \hat{B}_t.$$

Since the public debt can be freely chosen by the planner, it is equivalent for the planner to choose $\hat{B}_t$ rather than B_t.

The reformulated Ramsey program. We reformulate the the Ramsey program (70)–(76) using the variables $\tilde{a}_{i,t}$, W_t, R_t, $\tilde{A}_t$, $\hat{B}_t$ introduced in (62)–(64) and (68)–(69). The program can be expressed in post-tax prices R_t and W_t – taxes and pre-tax factor prices can be deduced from the allocation and the post-tax price definitions. The following proposition summarizes the reformulation of the Ramsey program.

Proposition 3 *The Ramsey program (70)–(76) can be rewritten as:*

$$\max_{\left(W_t, R_t, \tau_t, \hat{B}_t, \tilde{A}_t, K_t, L_t, (c_{i,t}, l_{i,t}, \tilde{a}_{i,t}, \nu_{i,t})_i\right)_{t\geq 0}} SWF_0, \tag{70}$$

$$G + W_t \int_i (y_{i,t} l_{i,t})^{1-\tau_t} \ell(di) + (R_t - 1)\tilde{A}_{t-1} + \hat{B}_{t-1} = F(K_{t-1}, L_t, s_t) + \hat{B}_t, \tag{71}$$

$$\text{for all } i \in \mathcal{I}\text{: } c_{i,t} + \tilde{a}_{i,t} = W_t (y_{i,t} l_{i,t})^{1-\tau_t} + R_t \tilde{a}_{i,t-1}, \tag{72}$$

$$\tilde{a}_{i,t} \geq -\bar{\tilde{a}}, \ \nu_{i,t}(\tilde{a}_{i,t} + \bar{\tilde{a}}) = 0, \ \nu_{i,t} \geq 0, \tag{73}$$

$$u'(c_{i,t}) = \beta \mathbb{E}_t \Big[R_{t+1} u'(c_{i,t+1}) \Big] + \nu_{i,t}, \tag{74}$$

$$v'(l_{i,t}) = (1-\tau_t) W_t y_{i,t} (y_{i,t} l_{i,t})^{-\tau_t} u'(c_{i,t}), \tag{75}$$

$$K_t + \hat{B}_t = \tilde{A}_t = \int_i \tilde{a}_{i,t} \ell(di), \ L_t = \int_i y_{i,t} l_{i,t} \ell(di). \tag{76}$$

C.2 Lagrangian and the FOCs of the Ramsey program

The Lagrangian associated to the Ramsey program (70)–(76) can be written as:

$$\begin{aligned}
\mathcal{L} = &\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \omega_{i,t} (u(c_{i,t}) - v(l_{i,t})) \ell(di) \\
&- \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i (\lambda_{c,i,t} - R_t \lambda_{c,i,t-1}) \, u'(c_{i,t}) \ell(di) \\
&- \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int_i \lambda_{l,i,t} \left(v'(l_{i,t}) - (1-\tau_t) W_t y_{i,t} (y_{i,t} l_{i,t})^{-\tau_t} u'(c_{i,t}) \right) \ell(di) \\
&- \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mu_t \left(G_t + (1-\delta)\hat{B}_{t-1} + (R_t - 1 + \delta) \int_i \tilde{a}_{i,t-1} \ell(di) + W_t \int_i (y_{i,t} l_{i,t})^{1-\tau_t} \ell(di) - Y_t - \hat{B}_t \right).
\end{aligned}$$

where the value of ν_{it} is given by the complementary slackneww conditions (73) and (74), and where we have

$$c_{i,t} = -\tilde{a}_{i,t} + R_t \tilde{a}_{i,t-1} + W_t (y_{i,t} l_{i,t})^{1-\tau_t}, \tag{77}$$

$$Y_t = \Big(\int_i \tilde{a}_{i,t-1} \ell(di) - \hat{B}_{t-1} \Big)^{\alpha} \Big(\int_i y_{i,t} l_{i,t} \ell\,(di) \Big)^{1-\alpha}. \tag{78}$$

As a consequence, the instruments are: $\tilde{a}_{i,t}$, $l_{i,t}$, W_t, R_t, τ_t, and $\hat{B}_t$. Using the two previous equations to substitute $c_{i,t}$ and Y_t, he program of the planner is

$$\max_{\left(W_t, R_t, \tau_t, \hat{B}_t, (l_{i,t}, \tilde{a}_{i,t})_i\right)_{t \geq 0}} \mathcal{L}$$

We now provide the first-order conditions of the planner and we present an alternative interpretation of the Lagragian in the next section.

FOC with respect to public debt $\hat{B}_t$.

$$\mu_t = \beta \mathbb{E}_t \left[(1 + \tilde{r}_{t+1}) \mu_{t+1} \right]. \tag{79}$$

FOC with respect to savings choices $\tilde{a}_{i,t}$. We define the marginal social value of liquidity for agent i at date t as:

$$\psi_{i,t} := \omega_{i,t} u'(c_{i,t}) - \Big(\lambda_{c,i,t} - R_t \lambda_{c,i,t-1} - \lambda_{l,i,t} (1 - \tau_t) W_t (y_{i,t})^{1-\tau_t} (l_{i,t})^{-\tau_t} \Big) u''(c_{i,t}), \tag{80}$$

and $\hat{\psi}_{i,t} := \psi_{i,t} - \mu_t$ as the marginal social value of liquidity net of the cost for planner's resources. We obtain using (79):

$$\hat{\psi}_{i,t} = \beta \mathbb{E}_t \left[R_{t+1} \hat{\psi}_{i,t+1} \right]. \tag{81}$$

FOC with respect to labor supply $l_{i,t}$. We define:

$$\psi_{l,i,t} := \omega_{i,t} v'(l_{i,t}) + \lambda_{l,i,t} v''(l_{i,t}).$$

The FOC with respect to labor supply $l_{i,t}$ is:

$$\psi_{l,i,t} = (1 - \tau_t) W_t (y_{i,t})^{1-\tau_t} (l_{i,t})^{-\tau_t} \hat{\psi}_{i,t} + \mu_t F_{L,t} y_{i,t} - (1 - \tau_t) W_t (y_{i,t})^{1-\tau_t} (l_{i,t})^{-\tau_t} \lambda_{l,i,t} \tau_t \frac{u'(c_{i,t})}{l_{i,t}}.$$

FOC with respect to the wage rate W_t.

$$0 = \int_j (y_{j,t} l_{j,t})^{1-\tau_t} \Big(\hat{\psi}_{j,t} + \lambda_{l,j,t} (1 - \tau_t) u'(c_{j,t}) / l_{j,t} \Big) \ell(dj).$$

FOC with respect to the interest rate R_t.

$$0 = \int_j (\hat{\psi}_{j,t} \tilde{a}^j_{t-1} + \lambda_{c,j,t-1} u'(c_{j,t})) \ell(dj). \tag{82}$$

FOC with respect to progressivity τ_t.

$$0 = \int_j (y_{j,t}l_{j,t})^{1-\tau_t}(\hat{\psi}_{j,t} + \lambda_{l,j,t}(1-\tau_t)(u'(c_{j,t})/l_{j,t}))\ln(y_{j,t}l_{j,t})\ell(dj)$$
$$+ \int_j \lambda_{l,j,t}(y_{j,t}l_{j,t})^{1-\tau_t}(u'(c_{j,t})/l_{j,t})\ell(dj).$$

C.3 Expression of Lagrangian using a public finance representation

The Lagrangian can be written as

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\mathcal{W}_t + \mu_t \mathcal{B}_t\right)$$

with

$$\mathcal{W}_t := \int_i \left(\omega_{i,t}(u(c_{i,t}) - v(l_{i,t})) - (\lambda_{c,i,t} - R_t\lambda_{c,i,t-1})\, u'(c_{i,t})\right.$$
$$\left. -\lambda_{l,i,t}\left(v'(l_{i,t}) - (1-\tau_t)W_t y_{i,t}(y_{i,t}l_{i,t})^{-\tau_t}u'(c_{i,t})\right)\right)\ell(di)$$
$$\mathcal{B}_t := Y_t - \hat{B}_t - G_t - (1-\delta)\hat{B}_{t-1} - (R_t - 1 + \delta)\int_i \tilde{a}_{i,t-1}\ell(di) - W_t\int_i (y_{i,t}l_{i,t})^{1-\tau_t}\ell(di)$$

The quantity $\mathcal{B}_t$ is the budget constraint of the government, whereas $\mathcal{W}_t$ is the aggregate welfare taking account the possible general equilibrium effects generated by each agent's choice, and captured by individual Lagrange multipliers $\lambda_{c,i,t}$ and $\lambda_{l,i,t}$. Note that if these multipliers were 0, then $\mathcal{W}_t$ would only be the weighted welfare.

Considering an instrument I_k in period k (I_k can be public debt, interest rate labor tax or its progressivity), the FOC of the Lagrangian with respect to I_k implies:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mu_t \frac{d\mathcal{B}_t}{dI_k} + \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int \frac{\partial \mathcal{W}_t}{\partial I_k}\ell(di) = -\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \int \frac{\partial \mathcal{W}_t}{\partial c_{i,t}}\frac{\partial c_{i,t}}{\partial I_k}\ell(di). \tag{83}$$

Since W_t, R_t, τ_t only affect the current value of welfare and budget constraint (for $I_t \in \{W_t, R_t, \tau_t\}$, $\frac{\partial \mathcal{W}_t}{\partial I_k} = \frac{\partial \mathcal{B}_t}{\partial I_k} = 0$ if $k \neq t$), we have for any $I_t \in \{W_t, R_t, \tau_t\}$:

$$\mu_t = \int \frac{\partial \mathcal{W}_t}{\partial c_{i,t}} \frac{-\frac{\partial c_{i,t}}{\partial I_t}}{\frac{d\mathcal{B}_t}{dI_t} + \frac{1}{\mu_t}\frac{\partial \mathcal{W}_t}{\partial I_t}}\ell(di). \tag{84}$$

It is then easy to compute the partial derivatives to check that we obtain the same expressions as in Section C.2. For instance, we have $\frac{\partial \mathcal{W}_t}{\partial c_{i,t}} = \psi_{i,t}$ and when the fiscal instrument is the post-tax rate R_t: $\frac{d\mathcal{B}_t}{dR_t} = -\int_i \tilde{a}_{i,t-1}\ell(di)$, $\frac{\partial \mathcal{W}_t}{\partial R_t} = \int_i R_t\lambda_{c,i,t-1}u'(c_{i,t})$, and $\frac{\partial c_{i,t}}{\partial R_t} = \tilde{a}_{i,t-1}$. Then, (84) becomes:

$$\mu_t = \int \psi_{i,t} \frac{-\tilde{a}_{i,t-1}}{-\int_i \tilde{a}_{i,t-1}\ell(di) + \frac{1}{\mu_t}\int_i R_t\lambda_{c,i,t-1}u'(c_{i,t})}\ell(di),$$

which can be written as $\int(\psi_{i,t} - \mu_t)\tilde{a}_{i,t-1}\ell(di) + \int_i R_t\lambda_{c,i,t-1}u'(c_{i,t}) = 0$, which is the FOC (82) with $\hat{\psi}_{i,t} = \psi_{i,t} - \mu_t$. The same derivations can be obtained for R_t, τ_t.

The instrument $\hat{B}_t$ is a fiscal instrument that affects current and future budget constraints but not welfare: $\frac{\partial \mathcal{W}_t}{\partial \hat{B}_k} = 0$, for all k and $\frac{\partial \mathcal{B}_t}{\partial \hat{B}_t} = -1$, $\frac{\partial \mathcal{B}_{t+1}}{\partial \hat{B}_t} = 1 + \tilde{r}_{t+1}$, and $\frac{\partial \mathcal{B}_\tau}{\partial \hat{B}_t} = 0$ if $\tau \notin$

$\{t, t+1\}$. As a consequence, the FOC (83) simplifies into0 $\mathbb{E}_0[\beta^t \mu_t \frac{d\mathcal{B}_t}{d\hat{B}_t} + \beta^{t+1}\mu_{t+1}\frac{d\mathcal{B}_{t+1}}{d\hat{B}_t}] = 0$, or $-\mu_t + \beta\mathbb{E}_t[\mu_{t+1}(1+\tilde{r}_{t+1})] = 0$, which is FOC (79).

D Truncating the model and identification of Pareto Weights

D.1 The truncated model

The key step of the aggregation consists of assigning the same wealth and allocation to all agents sharing the same idiosyncratic history over the recent past. The recent past is characterized by a number of periods, called the *truncation length* and denoted N; it is a parameter of the model. This N-period history will be referred to as a *truncated history*. For an history $y^t = \{\ldots, y^t_{t-N}, y^t_{t-N+1}, \ldots, y^t_{t-1}, y^t_t\}$, this corresponds to the N-length vector denoted $y^N := \{y^t_{t-N}, y^t_{t-N+1}, \ldots, y^t_{t-1}, y^t_t\}$. To sum up, we can represent the truncated history of an agent i whose idiosyncratic history is y^t as:

$$y^t = \{\underbrace{\ldots, y^t_{t-N-2}, y^t_{t-N-1}, y^t_{t-N}}_{\xi_{y^N}}, \underbrace{y^t_{t-N+1}, \ldots, y^t_{t-1}, y^t_t}_{=y^N}\}$$

where the parameter ξ_{y^N} captures the residual heterogeneity for the truncated history y^N, and y^t_{t-k} represents the idiosyncratic variable (at date t) k periods in the past. The method to compute the set of parameters $(\xi_{y^N})_{y^N}$ will be discussed further below. In what follows, we will discuss the various elements needed to apply the aggregation procedure.

First, we need to compute the measure of agents with the same history y^N. An agent with history $\tilde{y}^N$ at $t-1$ will have a different truncated history in period t depending on the realization of the idiosyncratic variable at date t. The probability to transit from truncated history $\tilde{y}^N$ to truncated history y^N will be denoted by $\Pi_{\tilde{y}^N y^N}$ (with $\sum_{y^N \mathcal{Y}^{\infty N}} \Pi_{\tilde{y}^N y^N} = 1$) and can be computed from the transition probabilities for the productivity process as:

$$\Pi_{\tilde{y}^N y^N} = 1_{y^N \succeq \tilde{y}^N} \Pi_{\tilde{y}^N_0 y^N_0} \geq 0,$$

where $1_{y^N \succeq \tilde{y}^N}$ is equal to 1 if y^N is a possible continuation of $\tilde{y}^N$, and 0 otherwise. With those elements, we can compute the share of agents with truncated history y^N as S_{t,y^N}. This element will be:

$$S_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} S_{t-1,\tilde{y}^N} \Pi_{\tilde{y}^N y^N}, \tag{85}$$

where the initial shares $(S_{-1,y^N})_{y^N \in \mathcal{Y}^N}$ are given with $\sum_{y^N \in \mathcal{Y}^N} S_{-1,y^N} = 1$.

The model aggregation then assigns to each truncated history the average choices (whether for consumption, savings, or labor supply) of the group of agents sharing the same truncated history. Let us consider a generic variable denoted by $X_t(y^t, s^t)$, and denote by X_{t,y^N} the average quantity of X assigned to truncated history y^N. Formally:

$$X_{t,y^N} = \frac{1}{S_{t,y^N}} \sum_{y^t \in \mathcal{Y}^{t+1} \big| (y^t_{t-N+1}, \ldots, y^t_{t-1}, y^t_t) = y^N} X_t(y^t, s^t) \mu_t(y^t), \tag{86}$$

where we remind that $\mu_t(y^t)$ is the measure of agents with history y^t. Definition (86) can be applied to consumption, savings, labor supply, and credit-constraint Lagrange multiplier. This leads to the quantities c_{t,y^N}, $\tilde{a}_{t,y^N}$, l_{t,y^N}, and ν_{t,y^N}, respectively. Note that applying (86) to beginning-of-period wealth involves accounting for the fact that agents with truncated history y^N at date t may come from various truncated histories at $t-1$. Specifically, this variable consists of the wealth of all agents with history y^N in period t but with any other possible history in $t-1$. Formally, the beginning-of-period wealth $\tilde{\tilde{a}}_{t,y^N}$ for truncated history y^N is:

$$\tilde{\tilde{a}}_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \tilde{a}_{t-1,\tilde{y}^N}. \tag{87}$$

We now define the various "ξs". First, we define $\xi^{u,0}_{y^N}$ as::

$$\sum_{y^t \in \mathcal{Y}^{t+1} \left| (y^t_{t-N+1},\ldots,y^t_{t-1},y^t_t)=y^N \right.} u(c_t(y^t)) = \xi^{u,0}_{y^N} u\Big(\sum_{y^t \in \mathcal{Y}^{t+1} \left| (y^t_{t-N+1},\ldots,y^t_{t-1},y^t_t)=y^N \right.} c_t(y^t)\Big),$$

or compactly as:

$$\sum_{y^t_i \in \mathcal{Y}^{t+1} | y^{t,N}_i = y^N} u(c_{i,t}) = \xi^{u,0}_{y^N} u(c_{t,y^N}). \tag{88}$$

The quantity $\xi^{u,0}_{y^N}$ reflects that aggregating utility levels is not equal to the utility of aggregated consumption. This comes from the combination of two reasons. First, there is heterogeneity of consumption among the population of agents having truncated history y^N, due their history prior to date $t-N$. Second, the utility function is not affine in general.

The same procedure applied to the other variables for the Ramsey problem (70)–(76) yields:

$$\sum_{y^t_i \in \mathcal{Y}^{t+1} | y^{t,N}_i = y^N} v(l_{i,t}) := \xi^{v,0}_{y^N} v(l_{t,y^N}), \tag{89}$$

$$\sum_{y^t_i \in \mathcal{Y}^{t+1} | y^{t,N}_i = y^N} u'(c_{i,t}) := \xi^{u,1}_{y^N} u'(c_{t,y^N}), \tag{90}$$

$$\sum_{y^t_i \in \mathcal{Y}^{t+1} | y^{t,N}_i = y^N} (y_{i,t} l_{i,t})^{1-\tau_t} := \xi^{y}_{y^N} (y^N_0 l_{t,y^N})^{1-\tau_t}. \tag{91}$$

We can now proceed with the aggregation of the full-fledged model. First, the aggregation of individual budget constraints (65) yields:

$$c_{t,y^N} + \tilde{a}_{t,y^N} = W_t \xi^{y}_{y^N} (l_{t,y^N} y^N_0)^{1-\tau_t} + R_t \tilde{\tilde{a}}_{t,y^N}, \text{ for } y^N \in \mathcal{Y}^{\infty N}. \tag{92}$$

The aggregation of Euler equations for consumption (74) and labor (75) yields:

$$\xi^{u,E}_{y^N} u'(c_{t,y^N}) = \beta \mathbb{E}_t \Big[R_{t+1} \sum_{\tilde{y}^N \in \mathcal{Y}^{\infty N}} \Pi_{t+1,y^N \tilde{y}^N} \xi^{u,E}_{\tilde{y}^N} u'(c_{t+1,\tilde{y}^N}) \Big] + \nu_{t,y^N}, \tag{93}$$

$$\xi^{v,1}_{y^N} v'(l_{t,y^N}) := (1-\tau_t) W_t \xi^{y}_{y^N} (l_{t,y^N} y^N_0)^{1-\tau_t} \xi^{u,1}_{y^N} (u'(c_{t,y^N})/l_{t,y^N}), \tag{94}$$

where the coefficients $(\xi^{u,E}_{y^N})_{y^N}$ for the consumption Euler equations ensure that the aggregate Euler equations yield Euler equations with aggregate consumption levels. In other words, the $(\xi^{u,E}_{y^N})_{y^N}$ are determined such that the aggregated consumption levels (for truncated histories) satisfy the consumption Euler equation (93). These coefficients are necessary because Euler equations involve non-linear marginal utilities. The same idea applies to the coefficients $(\xi^{v,1}_{y^N})_{y^N}$ for the FOC on labor.

Finally, market clearing conditions can be expressed as:

$$K_t + \hat{B}_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \tilde{a}_{t,y^N}, \quad L_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} y_{y^N} l_{t,y^N}. \tag{95}$$

Equations (92)–(95) exactly characterize the dynamics of the aggregated variables c_{t,y^N}, $\tilde{a}_{t,y^N}$, l_{t,y^N}, and ν_{t,y^N}, as well as aggregate quantities K_t, $\hat{B}_t$, and L_t.

Steady state and computation of the ξs. Steady-state allocations allow us to compute the parameters ξs as follows. We compute policy functions and wealth distribution of the Bewley model, as well as identify the set of credit-constrained histories, denoted $\mathcal{C}$. Aggregation equations (86) and (87) can then be used to aggregate (steady-state) allocations c_{y^N}, $\tilde{a}_{y^N}$, l_{y^N}, and ν_{y^N} . We then invert the consumption Euler equations (93) to deduce the preference parameters $(\xi^{u,E}_{y^N})_{y^N}$. The other ξs are computed explicitly by equations (88), (89), (90), (91), and (94).

The truncated model in the presence of aggregate shocks. We state two further assumptions that enable us to use the truncation method in the presence of aggregate shocks, resulting in the so-called truncated model.

Assumption B *We make the following two assumptions.*

1. *The preference parameters $(\xi_{y^N})_{y^N}$ remain constant and equal to their steady-state values.*
2. *The set of credit-constrained histories, denoted by $\mathcal{C} \subset \mathcal{Y}^{\infty N}$, is time-invariant.*

Two properties are finally worth mentioning. First, a straightforward consequence of the construction of the ξs is that the steady-state allocations of the initial and truncated models are identical. Second, as the truncation length N becomes increasingly long, truncated allocations (in the presence of aggregate shocks) can be shown to converge to those of the full-fledged equilibrium. Section 6 shows that from a quantitative standpoint, the ξs efficiently capture the heterogeneity within truncated histories, even when the truncation length remains limited.

D.2 Ramsey program

Program formulation. The finite state-space representation of the truncated model allows us to solve for the Ramsey program in the presence of aggregate shocks.[30] Let $(\omega_y)_{y \in \mathcal{Y}}$ denote the period weights associated with each productivity level. The Ramsey program in the truncated economy can be written as follows:

$$\max_{\left(W_t, R_t, \tilde{w}_t, \tilde{r}_t, \tau^c_t, \tau^K_t, \tau_t, \kappa_t, \hat{B}_t, G_t, K_t, L_t, (c_{t,y^N}, l_{t,y^N}, \tilde{a}_{t,y^N}, \nu_{t,y^N})_{y^N}\right)_{t \geq 0}} W_0, \tag{96}$$

[30]Our method involves deriving the FOCs of the truncated model, rather than truncating the FOCs of the full-fledged Ramsey model. This ensures numerical stability, as the truncated model is "well-defined" for the fiscal policy under consideration by construction.

where $W_0 := \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\sum_{y^N\in\mathcal{Y}^N} S_{t,y^N}\omega_{y^N}(\xi^{u,0}_{y^N}u(c_{t,y^N})-\xi^{v,0}_{y^N}v(l_{t,y^N})+u_G(G_t))\right]$ and subject to aggregate Euler equations (93) and (94), aggregate budget constraint (92), aggregate market clearing conditions (95), credit constraints $\tilde{a}_{t,y^N}\geq -\tilde{\bar{a}}$, as well as the governmental budget constraint (71), which is already present in the full-fledged Ramsey program.

First-order conditions. We define the net social value of liquidity of history y^N as in (80):

$$\begin{aligned}\hat{\psi}_{t,y^N} &= \omega_{y^N}\xi^{u,0}_{y^N}u'(c_{t,y^N})-\mu_t\\ &\quad-\left(\lambda_{c,t,y^N}\xi^{u,E}_{y^N}-R_t\tilde{\lambda}_{c,t,y^N}\xi^{u,E}_{y^N}-\lambda_{l,t,y^N}\xi^{y}_{y^N}(1-\tau_t)W_t(y_0^N)^{1-\tau_t}l_{t,y^N}^{-\tau_t}\xi^{u,1}_{y^N}\right)u''(c_{t,y^N}). \qquad (97)\end{aligned}$$

FOC with respect to $\tilde{a}_{t,y^N}$:

$$\hat{\psi}_{t,y^N} = \beta\mathbb{E}_t\Bigg[R_{t+1}\sum_{\tilde{y}^N\in\mathcal{Y}^{\infty N}}\Pi_{t,y^N\tilde{y}^N}\hat{\psi}_{t+1,\tilde{y}^N}\Bigg]\text{ if }\nu_{y^N}=0\text{ and }\lambda_{c,t,y^N}=0\text{ otherwise.} \qquad (98)$$

FOC with respect to l_{t,y^N}:

$$\begin{aligned}\frac{\omega_{y^N}\xi^{v,0}_{y^N}v'(l_{t,y^N})+\lambda_{l,t,y^N}\xi^{v,1}_{y^N}v''(l_{t,y^N})}{(1-\tau_t)W_t\xi^{y}_{y^N}(y_0^N)^{1-\tau_t}l_{t,y^N}^{-\tau_t}} &= \hat{\psi}_{t,y^N}-\lambda_{l,t,y^N}\tau_t\xi^{u,1}_{y^N}(u'(c_{t,y^N})/l_{t,y^N})\\ &\quad+\mu_t(1-\alpha)\frac{Y_t}{(1-\tau_t)W_t\xi^{y}_{y^N}(y_0^N)^{-\tau_t}l_{t,y^N}^{-\tau_t}L_t}. \qquad (99)\end{aligned}$$

FOC with respect to W_t:

$$\sum_{y^N\in\mathcal{Y}^{\infty N}} S_{t,y^N}\xi^{y}_{y^N}(l_{t,y^N}y_{y^N})^{1-\tau_t}\left(\hat{\psi}_{t,y^N}+\lambda_{l,t,y^N}(1-\tau_t)\xi^{u,1}_{y^N}(u'(c_{t,y^N})/l_{t,y^N})\right)=0. \qquad (100)$$

FOC with respect to R_t:

$$\sum_{y^N\in\mathcal{Y}^{\infty N}} S_{t,y^N}\left(\hat{\psi}_{t,y^N}\tilde{\tilde{a}}_{t,y^N}+\tilde{\lambda}_{c,t,y^N}\xi^{u,E}_{y^N}u'(c_{t,y^N})\right)=0. \qquad (101)$$

FOC with respect to τ_t:

$$\begin{aligned}&\sum_{y^N\in\mathcal{Y}^{\infty N}} S_{t,y^N}\left(\hat{\psi}_{t,y^N}+\lambda_{l,t,y^N}(1-\tau_t)\xi^{u,1}_{y^N}(u'(c_{t,y^N})/l_{t,y^N})\right)\ln\left(l_{t,y^N}y_{y^N}\right)\xi^{y}_{y^N}(l_{t,y^N}y_{y^N})^{1-\tau_t}\\ &=-\sum_{y^N\in\mathcal{Y}^{\infty N}} S_{t,y^N}\lambda_{l,t,y^N}\xi^{y}_{y^N}(l_{t,y^N}y_{y^N})^{1-\tau_t}\xi^{u,1}_{y^N}(u'(c_{t,y^N})/l_{t,y^N}). \qquad (102)\end{aligned}$$

FOC with respect to $\hat{B}_t$:

$$\mu_t=\beta\mathbb{E}\left[\mu_{t+1}\left(1+\alpha\frac{Y_{t+1}}{K_t}-\delta\right)\right]. \qquad (103)$$

We must furthermore have:

$$\tilde{\lambda}_{c,t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^{\infty N}} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \lambda_{c,t-1,\tilde{y}^N}, \tag{104}$$

$$\tilde{a}_{t,y^N} \geq 0 \text{ and } \tilde{\tilde{a}}_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^{\infty N}} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \tilde{a}_{t-1,\tilde{y}^N}. \tag{105}$$

D.3 Matrix expression

In this section, we provide closed-form formulas for preference multipliers ξs (Section D.1) and the weights ωs. We start with some 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.

The matrix representation consists in stacking together the equations characterizing the steady state, so as to provide a convenient matrix notation for solving the steady state. Truncated histories are simply indexed by y^N (the precise index does not matter as long as it remains the same). This also provides an efficient solution to compute the values for the coefficients (ξ_{y^N}) and (ω_{y^N}).

D.3.1 A closed-form formula for the ξs

Let $\boldsymbol{S} = (S_{y^N})_{y^N}$ be the N_{tot}-vector of steady-state history sizes (where N_{tot} is the number of truncated histories). Similarly, let $\boldsymbol{a}$, $\boldsymbol{c}$, $\boldsymbol{l}$, $\boldsymbol{\nu}$, $u'(\boldsymbol{c})$, $\boldsymbol{v'(l)}$ $\boldsymbol{u''(c)}$, $\boldsymbol{v''(l)}$ be the N_{tot}-vectors of end-of-period wealth, consumption, labor supply, Lagrange multipliers, marginal utilities, and derivatives of the marginal utility, respectively. These vectors are known from the steady-state equilibrium of the Bewley model. Each element is defined as the truncation of the relevant variable computed using equation (86). We also define by $\boldsymbol{y} = (y_0^N)_{y^N}$ the N_{tot}-vector of current productivity levels of truncated histories, and by $\boldsymbol{P}$ the diagonal matrix having 1 on the diagonal at y^N if and only if the history y^N is not credit constrained (i.e., $\nu_{y^N} = 0$), and 0 otherwise. Finally, $\boldsymbol{I}$ is the $(N_{tot} \times N_{tot})$-identity matrix, $\boldsymbol{\Pi}$ is the transition matrix across truncated histories.

Writing (85) , (92) and credit constraints at the steady state yield, respectively:

$$\boldsymbol{S} = \Pi \boldsymbol{S}, \tag{106}$$

$$\boldsymbol{S} \circ \boldsymbol{c} + \boldsymbol{S} \circ \tilde{\boldsymbol{a}} = R\Pi \left(\boldsymbol{S} \circ \tilde{\boldsymbol{a}}\right) + W\boldsymbol{S} \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}, \tag{107}$$

$$(\boldsymbol{I} - \boldsymbol{P})\tilde{\boldsymbol{a}} = \boldsymbol{0}_{N_{tot} \times 1}. \tag{108}$$

The Euler equation for consumption in (93) becomes:

$$\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}) = \beta R \boldsymbol{\Pi}^\top \left(\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c})\right) + \boldsymbol{\nu},$$

where the transpose matrix $\boldsymbol{\Pi}^\top$ implies expectations about next period histories. Equivalently:

$$\boldsymbol{D}_{u'(\boldsymbol{c})} \boldsymbol{\xi}^{u,E} = \beta R \boldsymbol{\Pi}^\top \boldsymbol{D}_{u'(\boldsymbol{c})} \boldsymbol{\xi}^{u,E} + \boldsymbol{\nu},$$

where $\boldsymbol{D}_{\boldsymbol{x}}$ stands for the diagonal matrix with the vector $\boldsymbol{x}$ on the diagonal. Finally:

$$\boldsymbol{\xi}^{u,E} = \left[\left(\boldsymbol{I} - \beta R \boldsymbol{\Pi}^{\top}\right) \boldsymbol{D}_{u'(\boldsymbol{c})}\right]^{-1} \boldsymbol{\nu}. \tag{109}$$

From the FOC on labor in (94), we obtain:

$$\boldsymbol{\xi}^{v,1} = (1-\tau) W (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \boldsymbol{\xi}^{y} \circ \boldsymbol{\xi}^{u,1} \circ u'(\boldsymbol{c})./(\boldsymbol{l} \circ v'(\boldsymbol{l})). \tag{110}$$

The equations (88)–(91) yield:

$$\boldsymbol{\xi}^{u,0} = \frac{\sum_{y^N \in \mathcal{Y}^{\infty N}} u(c_{i,t})}{u(c_{t,y^N})}, \quad \boldsymbol{\xi}^{u,1} = \frac{\sum_{y^N \in \mathcal{Y}^{\infty N}} u'(c_{i,t})}{u'(c_{t,y^N})}, \tag{111}$$

$$\boldsymbol{\xi}^{v,0} = \frac{\sum_{y^N \in \mathcal{Y}^{\infty N}} v(l_{i,t})}{v(l_{t,y^N})}, \quad \boldsymbol{\xi}^{y} = \frac{\sum_{y^N \in \mathcal{Y}^{\infty N}} (y_{i,t} l_{i,t})^{1-\tau}}{\left(y_0^N l_{t,y^N}\right)^{1-\tau}}. \tag{112}$$

Finally, we define the following variables:

D.3.2 Matrix expressions for the FOCs

We define the following variables: $\bar{\boldsymbol{\lambda}}_{\boldsymbol{l}} := \boldsymbol{S} \circ \boldsymbol{\lambda}_{\boldsymbol{l}}$, $\bar{\boldsymbol{\psi}} := \boldsymbol{S} \circ \hat{\boldsymbol{\psi}}$, $\bar{\boldsymbol{\Pi}} := \boldsymbol{S} \circ \boldsymbol{\Pi}^{\top} \circ (1/\boldsymbol{S})$, $\bar{\boldsymbol{\omega}} := \boldsymbol{S} \circ \boldsymbol{\omega}$, $\bar{\boldsymbol{\lambda}}_{\boldsymbol{c}} := \boldsymbol{S} \circ \boldsymbol{\lambda}_{\boldsymbol{c}}$, $\tilde{\boldsymbol{\xi}}^{u,1} := \boldsymbol{\xi}^{u,1}./\boldsymbol{l}$, $\tilde{\boldsymbol{\xi}}^{v,1} := \boldsymbol{\xi}^{v,1}./((1-\tau)W\boldsymbol{\xi}^{y} \circ \boldsymbol{y}^{1-\tau} \circ \boldsymbol{l}^{-\tau})$, and $\tilde{\boldsymbol{\xi}}^{v,0} := \boldsymbol{\xi}^{v,0}./((1-\tau)W\boldsymbol{\xi}^{y} \circ \boldsymbol{y}^{1-\tau} \circ \boldsymbol{l}^{-\tau})$. and notice that $\boldsymbol{S} \circ \tilde{\boldsymbol{\lambda}}_{\boldsymbol{c}} = \boldsymbol{\Pi}\bar{\boldsymbol{\lambda}}_{\boldsymbol{c}}$. With this notation, the FOCs (97)–(103) become:

$$\bar{\boldsymbol{\psi}} = \bar{\boldsymbol{\omega}} \circ \boldsymbol{\xi}^{u,0} \circ u'(\boldsymbol{c}) - \mu \boldsymbol{S} \tag{113}$$
$$-\left(\bar{\boldsymbol{\lambda}}_c \circ \boldsymbol{\xi}^{u,E} - R\boldsymbol{\Pi}\bar{\boldsymbol{\lambda}}_c \circ \boldsymbol{\xi}^{u,E} - (1-\tau)W\bar{\boldsymbol{\lambda}}_l \circ \boldsymbol{\xi}^{y} \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1}\right) \circ u''(\boldsymbol{c}),$$

$$\boldsymbol{P}\bar{\boldsymbol{\psi}} = \beta R \boldsymbol{P}\bar{\boldsymbol{\Pi}}\bar{\boldsymbol{\psi}}, \tag{114}$$

$$(\boldsymbol{I} - \boldsymbol{P})\bar{\boldsymbol{\lambda}}_c = 0, \tag{115}$$

$$\left(\boldsymbol{\xi}^{y} \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^{\top} \bar{\boldsymbol{\psi}} = -(1-\tau)\left(\boldsymbol{\xi}^{y} \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^{\top} \bar{\boldsymbol{\lambda}}_l, \tag{116}$$

$$\tilde{\boldsymbol{a}}^{\top}\bar{\boldsymbol{\psi}} = -\left(\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c})\right)^{\top} \Pi \bar{\boldsymbol{\lambda}}_c, \tag{117}$$

$$\bar{\boldsymbol{\omega}} \circ \tilde{\boldsymbol{\xi}}^{v,0} \circ v'(\boldsymbol{l}) + \bar{\boldsymbol{\lambda}}_l \circ \tilde{\boldsymbol{\xi}}^{v,1} \circ v''(\boldsymbol{l}) = \bar{\boldsymbol{\psi}} - \tau \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c}) \circ \bar{\boldsymbol{\lambda}}_l + \mu F_L \boldsymbol{S}./((1-\tau)W\boldsymbol{\xi}^{y} \circ \boldsymbol{y}^{-\tau} \circ \boldsymbol{l}^{-\tau}), \tag{118}$$

$$\left(\ln(\boldsymbol{y} \circ \boldsymbol{l}) \circ \boldsymbol{\xi}^{y} \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^{\top} \bar{\boldsymbol{\psi}} = -\left((\boldsymbol{1} + (1-\tau)\ln(\boldsymbol{y} \circ \boldsymbol{l})) \circ \boldsymbol{\xi}^{y} \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^{\top} \bar{\boldsymbol{\lambda}}_l. \tag{119}$$

D.3.3 Solving the system

Equation (118) yields:

$$\boldsymbol{D}_{\tilde{\boldsymbol{\xi}}^{v,1} \circ v''(\boldsymbol{l}) + \tau \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})} \bar{\boldsymbol{\lambda}}_l = \mu F_L \boldsymbol{S}./((1-\tau)W\boldsymbol{\xi}^{y} \circ \boldsymbol{y}^{-\tau} \circ \boldsymbol{l}^{-\tau}) + \bar{\boldsymbol{\psi}} - \boldsymbol{D}_{\tilde{\boldsymbol{\xi}}^{v,0} \circ v'(\boldsymbol{l})} \bar{\boldsymbol{\omega}},$$
$$\bar{\boldsymbol{\lambda}}_l = \boldsymbol{M}_0 \bar{\boldsymbol{\omega}} + \boldsymbol{M}_1 \bar{\boldsymbol{\psi}} + \mu \boldsymbol{V}_0. \tag{120}$$

with: $\boldsymbol{M}_0 := -\boldsymbol{M}_1 \boldsymbol{D}_{\tilde{\boldsymbol{\xi}}^{v,0} \circ v'(\boldsymbol{l})}$, $\boldsymbol{M}_1 := \boldsymbol{D}^{-1}_{\tilde{\boldsymbol{\xi}}^{v,1} \circ v''(\boldsymbol{l}) + \tau \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})}$, and $\boldsymbol{V}_0 := F_L \boldsymbol{M}_1 \boldsymbol{S}./((1-\tau)W\boldsymbol{\xi}^{y} \circ \boldsymbol{y}^{-\tau} \circ \boldsymbol{l}^{-\tau})$.

Equation (113) then implies:

$$\bar{\boldsymbol{\psi}} = \hat{\boldsymbol{M}}_0\bar{\boldsymbol{\omega}} + \hat{\boldsymbol{M}}_1\bar{\boldsymbol{\lambda}}_c + \hat{\boldsymbol{M}}_2\bar{\boldsymbol{\lambda}}_l - \mu\boldsymbol{S}. \tag{121}$$

with: $\hat{\boldsymbol{M}}_0 := \boldsymbol{D}_{\boldsymbol{\xi}^{u,0}\circ u'(\boldsymbol{c})}$, $\hat{\boldsymbol{M}}_1 := -\boldsymbol{D}_{\boldsymbol{\xi}^{u,E}\circ u''(\boldsymbol{c})}(\boldsymbol{I} - R\boldsymbol{\Pi})$, $\hat{\boldsymbol{M}}_2 := (1-\tau)W\boldsymbol{D}_{\boldsymbol{\xi}^{y}\circ(\boldsymbol{y}\circ\boldsymbol{l})^{1-\tau}\circ\tilde{\boldsymbol{\xi}}^{u,1}\circ u''(\boldsymbol{c})}$.

We obtain using (121) and (120):

$$\bar{\boldsymbol{\psi}} = \boldsymbol{M}_3\bar{\boldsymbol{\omega}} + \boldsymbol{M}_4\bar{\boldsymbol{\lambda}}_c + \mu\boldsymbol{V}_1, \tag{122}$$

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

Furthermore, equations (114), (115), and (122) imply:

$$\bar{\boldsymbol{\lambda}}_c = \boldsymbol{M}_5\bar{\boldsymbol{\omega}} + \mu\boldsymbol{V}_2, \tag{123}$$

where $\tilde{\boldsymbol{R}}_5 := -((\boldsymbol{I} - \boldsymbol{P}) + \boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})\boldsymbol{M}_4)^{-1}\boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})$, $\boldsymbol{M}_5 := \tilde{\boldsymbol{R}}_5\boldsymbol{M}_3$, and $\boldsymbol{V}_2 := \tilde{\boldsymbol{R}}_5\boldsymbol{V}_1$.

Substituting (122) and (123) into (117), we deduce:

$$\mu = -\boldsymbol{L}_0\bar{\omega}, \tag{124}$$

where $\boldsymbol{C}_1 := \tilde{\tilde{\boldsymbol{a}}}^\top(\boldsymbol{V}_1 + \boldsymbol{M}_4\boldsymbol{V}_2) + (\boldsymbol{\xi}^{u,E}\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}\boldsymbol{V}_2$ and $\boldsymbol{L}_0 := (\tilde{\boldsymbol{a}}^\top(\boldsymbol{M}_3 + \boldsymbol{M}_4\boldsymbol{M}_5) + (\boldsymbol{\xi}^{u,E}\circ u'(\boldsymbol{c}))^\top\boldsymbol{\Pi}\boldsymbol{M}_5)/\boldsymbol{C}_1$.

We deduce from (122) and (123):

$$\bar{\boldsymbol{\lambda}}_c = (\boldsymbol{M}_5 - \boldsymbol{V}_2\boldsymbol{L}_0)\bar{\boldsymbol{\omega}}, \tag{125}$$

$$\bar{\boldsymbol{\psi}} = \boldsymbol{M}_6\bar{\boldsymbol{\omega}}, \tag{126}$$

and from (120):

$$\bar{\boldsymbol{\lambda}}_l = \hat{\boldsymbol{M}}_6\bar{\boldsymbol{\omega}}. \tag{127}$$

We have defined $\hat{\boldsymbol{M}}_6 := \boldsymbol{M}_0 + \boldsymbol{M}_1\boldsymbol{M}_6 - \boldsymbol{V}_0\boldsymbol{L}_0$ and $\boldsymbol{M}_6 := \boldsymbol{M}_3 + \boldsymbol{M}_4(\boldsymbol{M}_5 - \boldsymbol{V}_2\boldsymbol{L}_0) - \boldsymbol{V}_1\boldsymbol{L}_0$.

Constructing the constraints. The constraint of equation (119) becomes after substituting the expressions (126) of $\bar{\boldsymbol{\psi}}$ and (127) of $\bar{\boldsymbol{\lambda}}_l$:

$$\tilde{\boldsymbol{L}}_1\bar{\boldsymbol{\omega}} = 0, \tag{128}$$

where:

$$\begin{aligned}\tilde{\boldsymbol{L}}_1 := &\left(\ln(\boldsymbol{y}\circ\boldsymbol{l})\circ\boldsymbol{\xi}^y\circ(\boldsymbol{y}\circ\boldsymbol{l})^{1-\tau}\right)^\top\boldsymbol{M}_6\\ &+\left((\boldsymbol{1} + (1-\tau)\ln(\boldsymbol{y}\circ\boldsymbol{l}))\circ\boldsymbol{\xi}^y\circ(\boldsymbol{y}\circ\boldsymbol{l})^{1-\tau}\circ\tilde{\boldsymbol{\xi}}^{u,1}\circ u'(\boldsymbol{c})\right)^\top\hat{\boldsymbol{M}}_6.\end{aligned}$$

The constraint (116) becomes after substituting the expressions (126) of $\bar{\boldsymbol{\psi}}$ and (127) of $\bar{\boldsymbol{\lambda}}_l$:

$$\tilde{\boldsymbol{L}}_2\bar{\boldsymbol{\omega}} = 0, \tag{129}$$

where:

$$\tilde{\boldsymbol{L}}_2 := \left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \boldsymbol{M}_6 + (1-\tau)\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^\top \hat{\boldsymbol{M}}_6.$$

The two constraints imposed on the history weights are $\tilde{\boldsymbol{L}}_1\bar{\boldsymbol{\omega}} = 0$ and $\tilde{\boldsymbol{L}}_2\bar{\boldsymbol{\omega}} = 0$. However, $\bar{\boldsymbol{\omega}}$ is a vector of length N_{tot}, while we care in Definition 2 about a vector $\boldsymbol{\omega}^Y = (\omega_y)_y$ of length Y. We define the $N_{tot} \times Y$-matrix $\boldsymbol{R}_0$ that maps a Y-vector into an N_{tot}-vector (where $\mathbf{1}_{S_y} \in \mathbb{R}^{S_y}$ is a S_y-vector of 1):

$$\boldsymbol{R}_0 := \begin{bmatrix} \mathbf{1}_{S_{y_1}} & 0 & 0 & 0 \\ 0 & \mathbf{1}_{S_{y_2}} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & S_{y_Y} \end{bmatrix},$$

and the $N_{tot} \times Y$-matrix $\boldsymbol{R}_1 := \boldsymbol{D}_{\boldsymbol{S}}\boldsymbol{R}_0$ that maps a Y-vector into an N_{tot}-vector, but where history sizes have been accounted for. To obtain dimensions compatible with other vectors and matrices, we define $\boldsymbol{\omega} = \boldsymbol{R}_0\boldsymbol{\omega}^Y$ and $\bar{\boldsymbol{\omega}} = \boldsymbol{R}_1\boldsymbol{\omega}^Y$.

In conclusion, the two constraints of Definition 2 on the weights $\boldsymbol{\omega}^Y = (\omega_y)_y$ are:

$$\boldsymbol{L}_1\boldsymbol{\omega}^Y = \boldsymbol{L}_2\boldsymbol{\omega}^Y, \tag{130}$$

with $\boldsymbol{L}_1 = \tilde{\boldsymbol{L}}_1\boldsymbol{R}_1$ and $\boldsymbol{L}_2 = \tilde{\boldsymbol{L}}_2\boldsymbol{R}_1$.

E Changes in the fiscal system

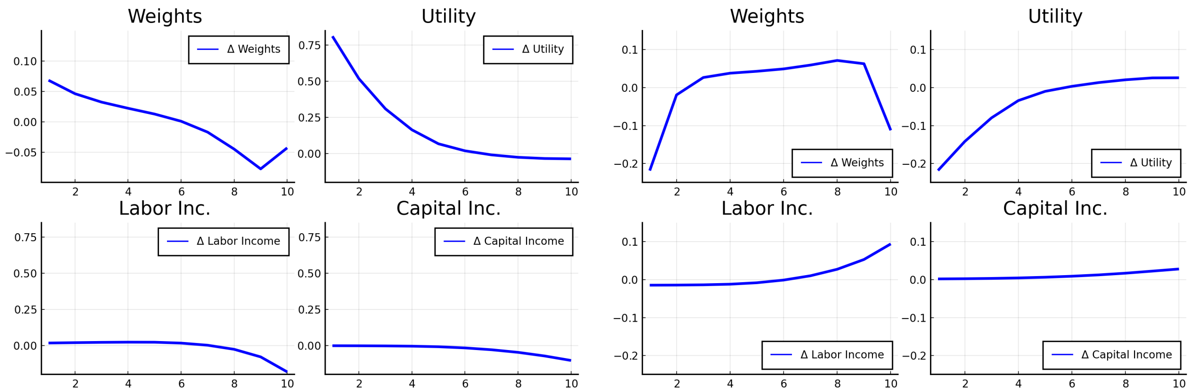
To identify the effect of the fiscal system in the indentification of weights, we now assume that fiscal systems are swapped between the two countries: France adopts the US fiscal system and the other way around. Table 6 summarizes the new fiscal system for each country.

	United States	**France**
τ_k	0.35	0.36
τ_c	0.18	0.05
κ	0.98	0.65
τ	0.23	0.16
B/Y	0.21	0.91

Table 6: New fiscal system for the United States and France in the current experiment.

We use our estimation strategy to compute the new SWF weights with the updated fiscal system of Table 6. We report in Figure 8 the differences implied by the new fiscal system compared to the benchmark for some key variables. These variables are SWF weights, utility level, labor supply, and capital income change. The results are averaged for each productivity level.

Considering the United States in panel (a), we observe that the change in the fiscal system increases the weights of low productivity agents and decreases the one of high-productivity. This results from low-productitivity agents benefiting from the new fiscal system, as can be see from the period utility, which decreases with productivity. We also observe that the new



(a) United States (b) France

Figure 8: Difference in weights between the US and French fiscal systems.

fiscal system makes both labor and capital income more progressive. Considering France in panel (b), we first observe the opposite variations. As can be seen from the decreasing utility, low-productivity agents suffer from the new fiscal system, which contributes to lower the SWF weights of low-productivity agents. The hump-shaped weights comes from high-productivity agents benefiting from the new fiscal system, which makes the labor and capital incomes less progressive.

F Robustness checks for SWF weights

We now relax two assumptions of our identification strategy: (i) the exact determination of weights by the constraints by imposing a parametric functional form; (ii) the weights depending only on the current productivity level.

F.1 Non-parametric weights

In the previous exercise we estimated parametric weights, where we imposed a functional relationship between weights and productivity to obtain an exact identification (see Definition 2). We consider here a different identification strategy to check the robustness of our results. We now estimate non-parametric weights, by choosing among all the weights verifying the constraints, those with the lowest variance.

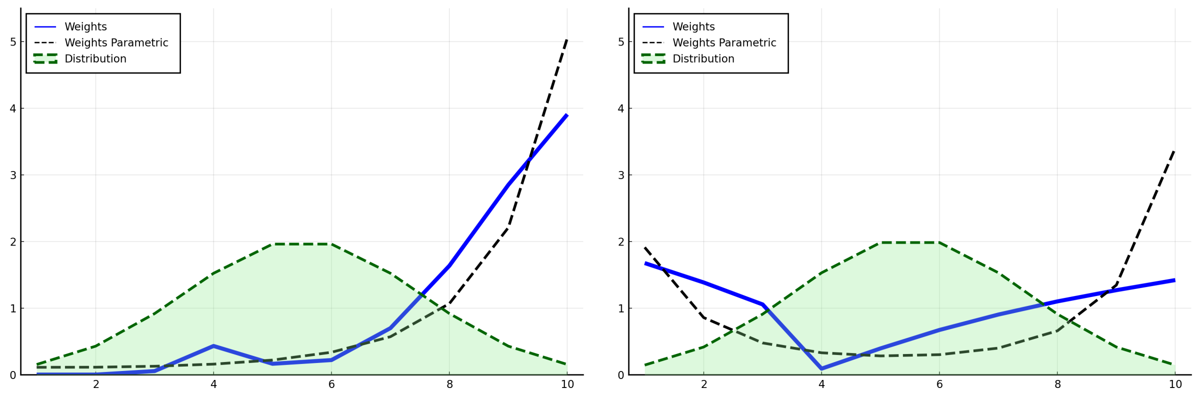
More precisely, as explained in the Section 4.4 (see equation (130)), the Ramsey FOCs impose two constraints: $\sum_{y\in\mathcal{Y}} L_{k,y}\omega_y = 0$, where $L_{k,y} \in \mathbb{R}$ ($k = 1, 2$ and $y \in \mathcal{Y}$). The variance-minimizing weights are characterized by the vector $(\hat{\omega}_y)_y$, solving the following program:

$$(\hat{\omega}_y)_y = \text{argmin}_{(\omega_y)_y} \sum_{y\in\mathcal{Y}} \pi_y(\omega_y - 1)^2, \tag{131}$$

$$\text{s.t. } 0 = \sum_{y\in\mathcal{Y}} L_{k,y}\omega_y \text{ for all } k = 1, 2, \tag{132}$$

$$1 = \sum_{y\in\mathcal{Y}} \pi_y\omega_y. \tag{133}$$

Figure 9 plots the non-parametric weights (blue solid lines) along the productivity dimension



(a) United States (b) France

Figure 9: Non-pametric weights (solid line) as a function of productivity levels for the US and France.

for the agents. We also report the parametric weights discussed in Section 6.2 (black dashed lines). both parametric and non-parametric weights are quite close to each other and exhibit a similar shape. The weights are increasing in the US and have a U-shape in France, with a high value of weights for low productivity agents. From this experience, we conclude that the shape of the weights is robust to the identification strategy, even if the value of weights for high productivity agents is not exactly identified in France.

F.2 Weights per truncated history

We relax here the assumption of the SWF weights depending solely on the current productivity level. We assume that weights possibly depend on the whole truncated history. We thus need to compute Y^N weights instead of Y. These weights are thus strongly under-identified. We use the same identification strategy as in for non-parametric weights in Section F.1. We select the minimal-variance weights verifying the constraints imposed by the Ramsey program. Formally, the weights $(\hat{\omega}_{y^N})_{y^N}$ are determined as follows:

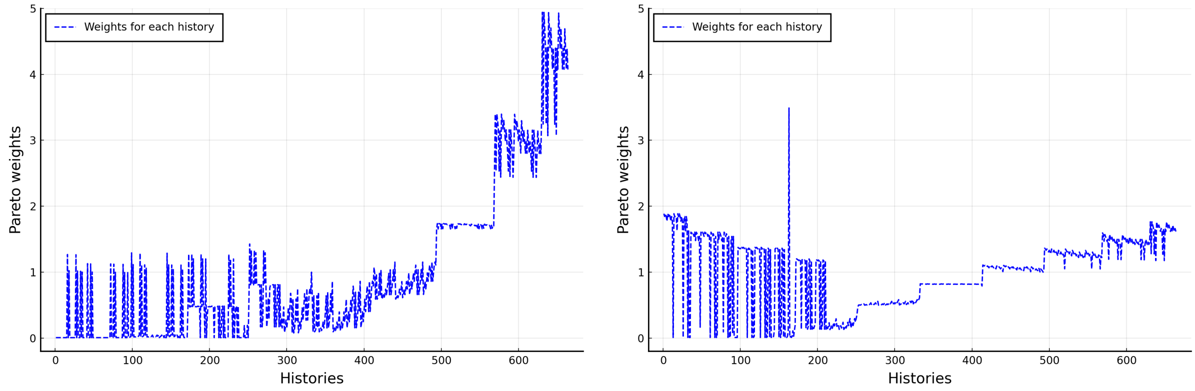
$$(\hat{\omega}_{y^N})_{y^N} = \text{argmin}_{(\omega_{y^N})_{y^N}} \sum_{y^N \in \mathcal{Y}^N} S_{y^N}(\omega_{y^N} - 1)^2, \tag{134}$$

$$\text{s.t. } 0 = \sum_{y^N \in \mathcal{Y}^N} \tilde{L}_{k,y^N} \omega_{y^N} \text{ for all } k = 1, 2, \tag{135}$$

$$1 = \sum_{y^N \in \mathcal{Y}^N} S_{y^N} \omega_{y^N}, \tag{136}$$

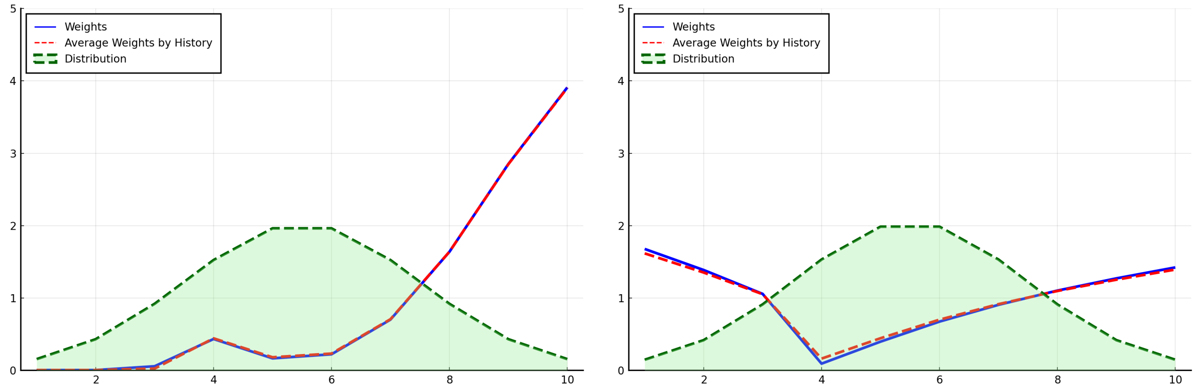
where $\tilde{L}_1$ and $\tilde{L}_2$ are defined in (128) and (128).

In Figure 10, we plot these history weights for the US in panel (a) and France in panel (b). We restrict to histories with a positive mass. To make them comparable with previous parametric weights we compute an average weight by summing the weights of truncated histories that have the same productivity level in the first period, and taking into account the size of each truncated history. This results into 10 weights, as for initial weights. The results are plotted in Figure 11, where we report the non-parametric Pareto weights as a solid blue line and the average history weights as a dashed red line (averaged over histories having the same



(a) United States (b) France

Figure 10: History weights for the US and France. Histories are arranged in the ascending order of the first-period productivity level.



(a) United States (b) France

Figure 11: Comparison between non-parametric weights and average history weights for the US and France.

current productivity level). We can observe that average history weights (red dashed line) closely approximate the non-parametric ones (blue line). Despite small differences, two methods imply very similar weights.