The Welfare of Nations: Do Social Preferences Matter for the Macroeconomy?

François Le Grand   Xavier Ragot   Diego Rodrigues*

PRELIMINARY, April 2022

Abstract

The levels of public debt, capital and labor taxes are very different across developed countries. To distinguish between the role of preferences and technologies in these differences, we derive a methodology to identify the Social Welfare Function (SWF) of a government, which is consistent with the empirical wealth and income distributions given the actual tax structure. We apply our methodology to identify the SWFs of France and the United States, using four fiscal instruments: consumption, capital and labor taxes, and public debt. The estimated SWFs of the two countries markedly differ from each other. The SWF for France gives a higher Pareto weight for individuals with lower income, whereas the United States SWF gives a higher social weight for agents at the top of income distribution. Finally, to assess the macroeconomic role of the SWF, we compute the fiscal system in the United States if they were to have the French SWF. We find that the SWFs affects the optimal steady-state fiscal system and equilibrium inequality, but surprisingly less the dynamics of the economy.

Keywords : Social Welfare Function, Inequality, Fiscal systems.

JEL codes : E61, E62, E32.

*LeGrand: emlyon business school and ETH Zurich; `legrand@em-lyon.com`. Ragot: SciencesPo, OFCE, and CNRS; `xavier.ragot@gmail.com`. Rodrigues: SciencesPo; `diego.desousarodrigues@sciencespo.fr`

1 Introduction

Countries differ widely in terms of their fiscal system: The levels of social transfers, labor and capital taxes are very different across OECD countries. For instance, the average mandatory levies in France is 40% of its GDP from 1995 to 2007, while it is 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, in addition to facing different trade-offs, preferences can be heterogeneous. Individual preferences over consumption and leisure, but also 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, which are observationally equivalent to alternative social preferences.

The goal of this paper is to identify Social Welfare Function (SWFs) from observed fiscal policy and redistributive outcome in both France and the US. This *positive* approach to *normative* economies is thus agnostic about the most relevant SWFs. We do not make any claim regarding the possibly optimality of SWFs. If economists or citizens have their own preferred SWFs, this estimation provides a measure of how different is the actual SWFs to their preferred ones. We are not the first to estimate SWFs, which is done in the literature studying the so-called Inverse Optimal Taxation Problem (see Heathcote & Tsujiyama 2021 for a recent paper). Our contribution is to provide a methodology to estimate SWF from quantitative heterogeneous-agent models while accounting for the general equilibrium effects generated by capital accumulation and for a rich set of fiscal instruments, such as capital and consumption taxes, non-linear labor tax, and public debt. We first show that an heterogeneous-agent model can accurately reproduce both France and US fiscal policies and equilibrium inequality. We then estimate the French and US SWFs, which would rationalize these choices of instruments. Our procedure relies on two main identifying assumptions.

First, we assume that the fiscal choices result from the optimal choices of a benevolent planner, which maximizes intertemporal welfare in an heterogeneous-agent model. The planner is endowed with her own SWFs, and understands distortions and the general-equilibrium effects of all fiscal instruments. The rationale for this strong identification strategy is the following. There is a vast literature on social choice studying the various political systems and their potential biases towards some clientele or some part of the population. Our identification strategy will attribute to the estimated SWFs all these biases in political systems. At an abstract level, one could argue that the political systems themselves are the outcome of choices. The estimated weights will thus absorb implicit or explicit choices embedded in political institutions, as is done in the inverse optimal taxation problem.

Second, as the number of instruments is finite and there is possibly a vast set of SWFs which could rationalize the fiscal outcomes, our second identification assumption is to select among possible SWFs the one which is the closest to the Utilitarian SWF – which attributes the same weight to all agents. This conservative assumption is sufficient to identify the SWF. We check the non-parametric Pareto weights are consistent with parametric weights relying on a specification à la Heathcote & Tsujiyama (2021).

We apply this methodology to identify the SWFs of France and the US, using first their fiscal system before the financial crisis. This choice is motivated by the fact that these two

OECD countries are very different in terms of fiscal systems. We choose the period from 1995 to 2007 because it does not cover the financial crisis and obviously the Covid-19 crisis, which have generated important (transitory) changes in fiscal structures.

This estimation generates two sets of results. First, the SWFs in France and in the US are very different from each other. The US SWF is mostly increasing with income and puts the largest weight to rich and high-income agents. The weights for the middle-class agents are lower and somewhat similar, and low income agents have the lowest weight. The French SWF is different. It puts the highest weight to low income agents, and the weight is decreasing with income, and increasing again for high income agents.

Our second result concerns the importance of these different SWFs for the dynamic of macroeconomics, that we investigate through the response to aggregate shock. We first, compute the optimal fiscal system considering the US economy where the planner is endowed with the French SWF. The first outcome is a very different steady-state fiscal system, which confirms the sensitivity of the optimal steady-state policy to the SWF. The second outcome is that the dynamic response of the US economy is surprisingly similar when the planner is endowed with US or French SWF. We thus conclude that social preferences are crucial for the design of tax system, but much less so for the analysis of business cycle properties.

Finally, to perform these exercises, we use the sequential representation of heterogeneous-agent models, together with a truncation procedure to obtain a finite state space (as in LeGrand & Ragot 2022). We then use the Lagrangian approach developed in Marcet & Marimon (2019) to derive first-order conditions of the Ramsey problem with commitment.

Related literature. Our paper is related to two streams of literature. First we contribute to the optimal fiscal policy literature, and to Inverse Optimal Taxation Problem, which estimate SWFs from actual fiscal policies. Most of this literature estimates the SWF, using only labor tax in economies without capital (Bargain & Keane 2010, Bourguignon & Amadeo 2015). Chang et al. (2018) use an heterogeneous-agent model with capital to estimate a parameter of the SWF, which is the concern for inequality. They focus on a single fiscal instrument. As we derive the first-order conditions of the Ramsey problem, we can estimate the whole set of Pareto weights using four fiscal instruments. Heathcote & Tsujiyama (2021) also estimate the SWF with one degree of freedom. Their focus is slightly different from ours since their framework is static, but features private information à la Mirrlees.

Second, this paper contributes to the new literature on optimal policies in heterogeneous-agent models. Some papers have studied the effects of given fiscal experiments in heterogeneous-agent frameworks, such as Heathcote (2005), who considers aggregate shocks, and Kaplan & Violante (2014), who consider a fiscal transfer. Heathcote & Perri (2017) analyze equilibrium multiplicity in an economy without capital. In economies without aggregate shocks, Aiyagari (1995) shows that the steady-state capital tax can be non-negative. Aiyagari & 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 recent papers solves a Ramsey problem to obtain the steady-state fiscal policy and level of public debt. Dyrda & Pedroni (2018) maximize welfare over all transition paths to determine optimal tax system. Açikgöz et al. (2018) use the first-order conditions of the Ramsey

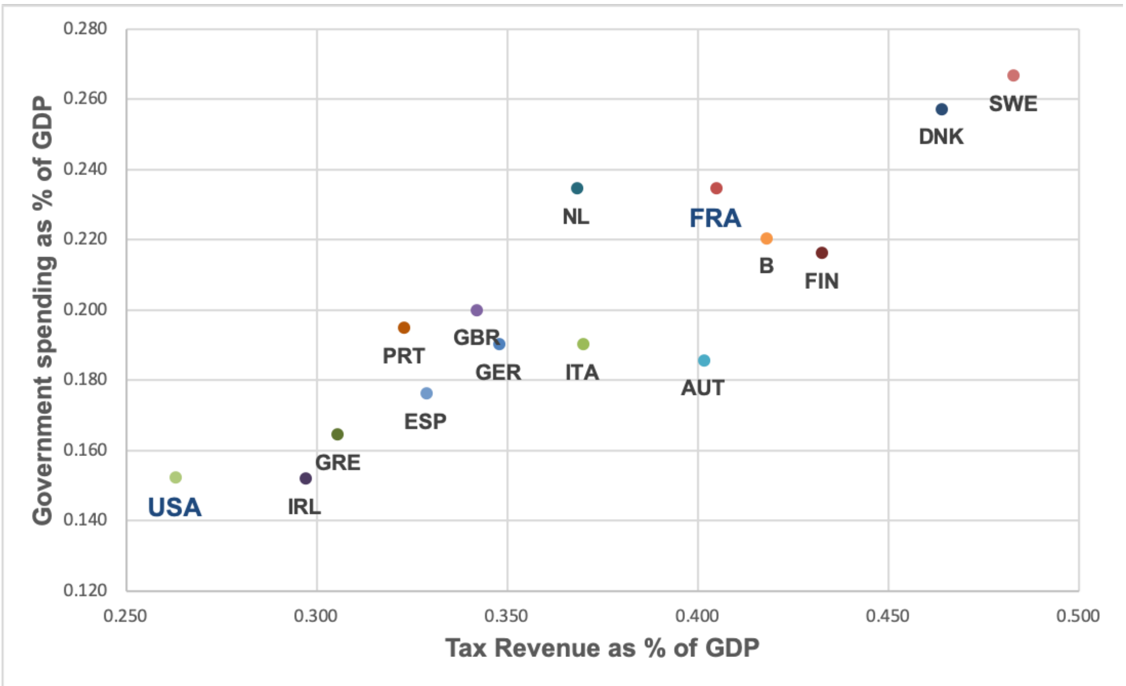


Figure 1: Government spending and tax revenues average from 1995 to 2007 (as a share of GDP). Source Trabandt & Uhlig (2011).

problem to derive the optimal tax system. These two papers consider the Utilitarian SWF.

Finally, there is a vast literature on social preferences, considering fairness principles such as desert or responsibility. In this literature, we follow the weighted-utilitarianism approach, which is flexible enough to incorporate these concerns (Fleurbaey & Maniquet 2018). This approach although simpler than the recent contribution of Saez & Stantcheva (2016) allows us to estimate SWF in general equilibrium.

The paper proceeds as follows. In Section 2 we present motivating evidence. Sections 3 to 5 explain the methodology. Finally, the quantitative investigation is presented 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. As explained in the introduction, these two countries have the particularity to be among the most different OECD countries in terms of total taxation. On the one hand, France is one of the countries with the highest mandatory levy, while the US is one of the countries with the lowest levy. This is confirmed in Figure 1, which reports the government spending on final goods as a share of GDP and tax revenues as a share of GDP for the two countries. A first observation from Figure 1 is that France and the US drastically differ with respect to the size of their government. Both governmental spending and tax revenues are significantly higher in France than in the US. Second the gap between tax revenues and governmental spending is larger in France than in the US. This reflects that within-country redistribution – measured as the difference between tax revenues and governmental 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 – that we will use as a benchmark for calibrating our initial stationary equilibrium. We use the results of Trabandt & Uhlig (2011), who provide estimates for the period 1995-2007. Results for France and the US are gathered in Table 1, where some elements regarding inequalities are also reported.

	Tot	τ^K	τ^L	τ^c	B	G	Gini bef.	Gini aft.	Gini wealth
France	40	35	46	18	60	24	.48	.28	.68
United States	26	36	28	5	63	15	.48	.40	.77

Table 1: Summary of fiscal systems and inequalities in the US and and in France.

The first column reports the total mandatory levies, as a share of GDP, for the two countries. Following the literature, we decompose 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 reports the implicit capital tax, which is the tax receipts on capital income divided by the capital tax. The third column is the same statistic for the labor tax , while 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 to each other in both countries, the labor and consumption taxes are very different. For labor tax, this difference partly stems from the financing of the French welfare system (that covers public pension, unemployment benefits, health care, and family allowances). Indeed, 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, even though this high value is comparable to the ones in other European countries.

The tax revenues are used to finance public spending. Public spending include public consumption and public investment and is approximately one-fourth higher in France than in the US. This difference is partly explained by larger spending in public infrastructure in France.

The results in Table 1 considers linear tax for labor, however, progressive labor income tax is commonly used as a fiscal instrument for countries and the current evidence suggests that the progressivity rate differs between France and the US. The problem is that comparing the progressivity of labor income tax across countries is not an easy task because of the complex tax schedules and deductions which are specific to each country. One tractable way to compare the progressivity is to use a parametric form for the tax function. To compute this value we use the log-linear tax function currently used in the literature (e.g. Benabou (2002) and Heathcote et al. (2017)).

$$\begin{aligned}\text{Tax:} \qquad & T(y_i) = y_i - \kappa y_i^{1-\tau}. \\ \text{Disposable income:} \qquad & D(y_i) = \kappa y_i^{1-\tau}. \\ & \log D(y_i) = \log \kappa + (1-\tau) \log y_i,\end{aligned}$$

where y_i denotes the labor income of the country i, the parameter τ reflect the level of progressivity, and κ the average level of taxation. Notice that the higher the τ, more progressive is the tax system. Using the Luxembourg Income Study (LIS) database for France and the US for 2005 we estimate the tax progressivity for the labor income ($\hat{\tau}$). In order to estimate this parameter we restrict our attention to the head of the households and their spouses with age between 25 and 60 and who were employed. We define the *labour income* as being the sum of wage income, self-employment income, and private transfers. Using the estimate of Mendoza et al. (1994) we can deduce from the capital income the value of capital income tax and we subtract from the total income tax this last variable. This allow us to obtain an estimate of the labor income tax and we can proceed by defining the *disposable income* as being the *labor income* minus the *labor income tax*. Finally we regress the log of disposable income in the log of labor income.

Table 2 reports our estimates for France and the US. One can notice that France has a much more progressive labor tax than the US. Our estimate of progressivity for US is 0.16, which is closely related to the value used in the literature. Our value is lower than the value of 0.181 estimated by Heathcote et al. (2017) because we just estimate the progressivity of the labor income and did not consider an estimate of the progressivity for the labor and capital income together.

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

Table 2: Estimate of the progressivity of the labor income tax in the US and and in France for 2005 using the LIS database.

All this heterogeneity in redistribution is also reflected in the evolution of income inequality before and after taxation. We proxy the income inequality by the average Gini index between 1995 and 2007 (included), as reported in the OECD Income Inequality Database.[1] Note that the Gini indices barely vary in 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. In consequence, this contributes to increase the before-tax inequalities. However, the after-tax Gini indices are very different in the two countries, which is a consequence in terms of inequalities of the high transfers to households in France. While redistribution diminishes Gini for income by less than 10 points in the US, the reduction is twice larger in France and amounts to 20 points.

The before-last element we report is the public debt-to-GDP ratio, that appears to be comparable in France and in the US at the value of 60% approximately.

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 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.[2] As is standard, the wealth inequalities in each country are higher for wealth than for income. The wealth Gini index in each country is approximately 30 points higher for wealth than for post-tax income. The comparison between the US and France yields a similar result as the income Gini index did. It indeed confirms that inequalities – for wealth here – are sharper in the US than in France.

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

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

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

3 The model

Time is discrete, indexed by $t \geq 0$. The economy is populated by a continuum of ex-ante identical agents. The population of size 1 is distributed on a segment J following a non-atomic measure ℓ: $\ell(J) = 1$. We follow Green (1994) and assume that the law of large number holds.

3.1 Risk structure

Aggregate risk. The aggregate risk is represented by a probability space $(\mathcal{S}^\infty, \mathcal{F}, \mathbb{P})$. In any period t, the aggregate state, denoted s_t, takes values in the state space $\mathcal{S} \subset \mathbb{R}^+$ and follows a first-order Markov process. The history of aggregate shocks up to time t is denoted by $s^t = \{s_0, \ldots, s_t\} \in \mathcal{S}^{t+1}$. The period-0 probability density function of any history s^t is denoted by $m_t(s^t)$.

Idiosyncratic risk. At the beginning of each period, agents face an uninsurable idiosyncratic labor productivity shock y_t that can take Y distinct values in the set $\mathcal{Y} \subset \mathbb{R}_+$. The productivity shock y_t follows a first-order Markov process with transition matrix $M_t \in [0,1]^{Y \times Y}$. The history of idiosyncratic shocks up to date t is denoted by $y^t = \{y_0, \ldots, y_t\} \in \mathcal{Y}^{t+1}$. We also denote by θ_t the measure of date-t idiosyncratic histories, that can be deduced from transition probabilities. More precisely $\theta_t(y^t)$ corresponds to the share of agents with history y^t at date t.

Initial wealth. In period 0, agents draw initial productivity and asset (y_0, a_{-1}) out of an initial distribution $\mathbf{\Lambda}_0(y,a)$ defined over the Borel sets of $\mathcal{Y} \times \mathbb{R}$, with $\mathbf{\Lambda}_0(y,a) : \mathcal{Y} \times \mathbb{R} \to \mathbb{R}^+$. This allows us to model an economy starting from an arbitrary distribution, including the steady-state distribution.

Remark 1 (Simplifying Notation)**.** *If an agent has an idiosyncratic history y_i^t, and initial wealth a_{-1}^i at period t, where the aggregate history is s^t, we will then denote the realization in state (y_i^t, a_{-1}^i, s^t) of any random variable $X_t : \mathcal{Y}^{t+1} \times \mathbb{R} \times \mathcal{S}^{t+1} \to \mathbb{R}$ simply by X_t^i.*

As a consequence, the aggregation of the variable X_t at period t over all agents will be written as $\int_i X_t^i \ell(di)$, instead of the more involved explicit notation where we use the sequential representation and integrate over initial states (of measure Λ) and idiosyncratic histories (of measure θ), considering the realization of the aggregate variable s^t:

$$\int_{a_{-1}} \sum_{y^t \in \mathcal{Y}^{t+1}} \theta_t\left(y^t\right) X_t\left(y^t, a_{-1}, s^t\right) \mathbf{\Lambda}\left(y_0, da_{-1}\right).$$

3.2 Production and government

Production. In any period t, a production technology with constant returns to scale (CRS) transforms capital K_{t-1} and labor L_t into $F(K_{t-1}, L_t, s_t)$ units of output. The production function is smooth in K and L and satisfies the standard Inada conditions. Furthermore, it features constant-return to scale and is homogeneous of degree 1. Capital must be installed one period before production, and the state of the world may potentially affect productivity through a TFP shock. 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 $t \geq 1$:

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

Government. A benevolent government has to choose a path stream of public spending, denoted by $G_t \equiv G_t(s^t)$ and to finance it. Several instruments are available. First, the government can levy one-period public debt B_t. We assume the existence of an enforcement technology that makes the public debt default-free. The public debt and capital are assumed to be perfectly substitute and to payoff the same pre-tax interest rate $\tilde{r}_t$. This assumption is akin to the existence of a risk-neutral fund (see Gornemann et al. 2016 among others) holding all interest-bearing assets (i.e., capital and public debt) and selling its shares to agents.

Second, the government can raise a number of distortionary taxes, which concern consumption, labor income, and capital revenues. The tax on labor income, denoted by $\mathcal{T}_t(\tilde{w}yl)$ for a labor income $\tilde{w}yl$, is assumed to be non-linear, and possibly time-varying. We follow Heathcote et al. (2017) (henceforth, HSV) and assume that $\mathcal{T}_t$ is defined as follows:

$$\mathcal{T}_t(\tilde{w}yl) := \tilde{w}yl - \kappa_t(\tilde{w}yl)^{1-\tau_t}, \tag{2}$$

where κ captures the level of labor taxation and τ the progressivity. Both parameters are assumed to be time-varying and will be planner's instruments. When $\tau_t = 0$, labor tax is linear with 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 (2), combined with a linear capital tax, allows one to realistically reproduce the actual US system and its progressivity (see Ferriere & Navarro 2020).[3]

Consumption and capital taxes are linear and are denoted by τ_t^c, and τ_t^K at date t. All of these taxes are proportional taxes and imply a total governmental revenue equal to $\tau_t^c C_t + \int_i \mathcal{T}_t(\tilde{w}_t y_t^i l_t^i)\ell(di) + \tau_t^K \tilde{r}_t(K_{t-1} + B_{t-1})$, where C_t is the aggregate consumption and $K_{t-1} + B_{t-1}$ is aggregate savings in period t (i.e., capital plus public debt). From now on the aggregate savings will be denoted by A_{t-1} such that $\tau_t^K \tilde{r}_t(K_{t-1} + B_{t-1})$ are capital revenues in period t.

[3] The literature uses either the combination of a linear tax and of a lump-sum transfer (e.g., Dyrda & Pedroni 2018, Açikgöz et al. 2018) or the HSV structure. Heathcote & Tsujiyama (2021) show that the HSV structure is quantitatively more relevant. Opting for the HSV tax structure enables us to discuss the dynamics of optimal tax progressivity, following a public spending shock.

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

$$G_t + (1+\tilde{r}_t)B_{t-1} = \tau_t^c C_t + \int_i \mathcal{T}_t(\tilde{w}_t y_t^i l_t^i)\ell(di) + \tau_t^K \tilde{r}_t A_{t-1} + B_t. \tag{3}$$

We define post-tax rates r_t and w_t, as follows:[4]

$$r_t := (1-\tau_t^K)\tilde{r}_t, \quad w_t := \kappa_t(\tilde{w}_t)^{1-\tau_t}. \tag{4}$$

Using the property of constant-return-to-scale for F and the definition of post-tax rates (4), the governmental budget constraint can be written as:

$$G_t + (1+r_t)B_{t-1} + w_t \int_i (y_t^i l_t^i)^{1-\tau_t}\ell(di) + r_t K_{t-1} = \tau_t^c C_t + F(K_{t-1}, L_t, s_t) + B_t. \tag{5}$$

3.3 Households preferences and program

Preferences. Agents are expected-utility maximizers with a time-separable utility function, whose constant discount factor is denoted by $\beta \in (0,1)$. Their period utility function $U(c,l,G)$ over private consumption c, labor supply l, and public good consumption G is assumed to be separable so as to avoid the quantitative pitfalls of the GHH utility function (Auclert et al. 2021):[5]

$$U(c,l,G) := u(c) - v(l) + u_G(G). \tag{6}$$

The functions $u, u_G : \mathbb{R}_+ \to \mathbb{R}$ are twice continuously derivable, strictly increasing, and strictly concave, with $u'(0) = u_G'(0) = \infty$, while $v : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously derivable, strictly increasing, and strictly convex, with $v'(0) = 0$.

Agents' program. We consider an agent $i \in \mathcal{I}$. Her resources are made of labor income and saving payoffs. The post-tax labor income of an agent with productivity y_t^i and supplying the labor effort l_t^i amounts to $\tilde{w}_t y_t^i l_t^i - \mathcal{T}_t(\tilde{w}_t y_t^i l_t^i) = w_t(y_t^i l_t^i)^{1-\tau_t}$. Since public debt and capital shares are perfectly substitutes and pay the same post-tax interest rate r_t, savings payoffs are equal to $(1+r_t)a_{t-1}^i$ where a_{t-1}^i is the end-of-period-$t-1$ saving of agent i. The agent uses these resources to save and to consume. Consumption is taxed with rate τ_t^c. Since the agent considers the public good path $(G_t)_{t\geq 0}$ as exogenous, her program can formally be expressed as follows:

$$\max_{\{c_t^i, l_t^i, a_t^i\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(c_t^i) - v(l_t^i) + u_G(G_t)\right), \tag{7}$$

$$(1+\tau_t^c)c_t^i + a_t^i \leq w_t(y_t^i l_t^i)^{1-\tau_t} + (1+r_t)a_{t-1}^i, \tag{8}$$

$$a_t^i \geq -\bar{a}, c_t^i > 0, l_t^i > 0, \tag{9}$$

where $\mathbb{E}_0$ an expectation operator (with respect to aggregate and idiosyncratic risks), and where the initial state (y_0^i, a_{-1}^i, s_0) is given. At date 0, the agent decides her consumption $(c_t^i)_{t\geq 0}$, her labor supply $(l_t^i)_{t\geq 0}$, and her saving plans $(a_t^i)_{t\geq 0}$ that maximize her intertemporal utility of equation (7), subject to a budget constraint (8) and the previous borrowing limit (9). These

[4] To simplify the notation, r and w denote the *after-tax* wage and interest rates.

[5] All our results can be derived with a general utility function $U(c,l,G)$. A GHH utility function slightly simplifies the algebra, especially when deriving the Ramsey problem in Section 4.

decisions are functions of the initial state (y_0^i, a_{-1}^i, s_0), of the history of idiosyncratic shocks y_i^t and of the history of shocks s^t. Thus, there exist sequence of functions defined over $\mathcal{Y}^{t+1} \times \mathbb{R} \times \mathcal{S}^{t+1}$ and denoted by $(c_t, l_t, a_t)_{t\geq 0}$, such that the agent's optimal decision can be written as:

$$c_t^i = c_t(y_i^t, a_{-1}^i, s^t),$$
$$l_t^i = l_t(y_i^t, a_{-1}^i, s^t),$$
$$a_t^i = a_t(y_i^t, a_{-1}^i, s^t).$$

In what follows we simplify the notation and keep the $i-$ index following the notation of Remark 1.[6]

The first-order conditions (FOCs) associated to the agent's program (7)–(9) can be written as follows:

$$u'(c_t^i) = \beta \mathbb{E}_t \left[(1 + r_{t+1}) \frac{1 + \tau_t^c}{1 + \tau_{t+1}^c} u'(c_{t+1}^i) \right] + \nu_t^i, \tag{10}$$

$$v'(l_t^i) = \frac{(1 - \tau_t)}{(1 + \tau_t^c)} w_t y_t^i (y_t^i l_t^i)^{-\tau_t} u'(c_t^i), \tag{11}$$

where the quantity $\beta^t \nu_t^i$ denotes the Lagrange multiplier on agent i's credit constraint.

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

$$\int_i a_t^i \ell(di) = A_t = B_t + K_t, \quad \int_i y_t^i l_t^i \ell\,(di) = L_t. \tag{12}$$

For the sake of simplicity we formulate market clearing by integration over agents i. Equivalently, it could be possible to formulate the market clearing by integrating over initial states and idiosyncratic histories.[7]

The economy-wide resource constraint can be written as:

$$G_t + C_t + K_t - K_{t-1} = F(K_{t-1}, L_t, s_t), \tag{13}$$

where $C_t = \int_i c_t^i \ell(di)$ is the aggregate consumption.

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

Definition 1 (Sequential equilibrium). *A sequential competitive equilibrium is a collection of individual allocations $(c_t^i, l_t^i, a_t^i, \nu_t^i)_{t\geq 0, i\in\mathcal{I}}$, 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 $(\tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t, G_t)_{t\geq 0}$, such that, for an initial productivity and wealth distribution $(y_0^i, a_{-1}^i)_{i\in\mathcal{I}}$, and for initial values of capital stock and public debt verifying $K_{-1} + B_{-1} = \int_i a_{-1}^i \ell(di)$ and of the aggregate shock s_0, we have:*

1. *given prices, the functions $(c_t^i, l_t^i, a_t^i, \nu_t^i)_{t\geq 0, i\in\mathcal{I}}$ solve the agent's optimization program in*

[6] The existence of those functions can be found in Açikgöz (2016), Cheridito & Sagredo (2016), and Miao (2006).

[7] Using the sequential representation $\int_i a_t^i \ell(di)$ can be written as $\int_{a_{-1}} \sum_{y^t \in \mathcal{Y}^{t+1}} \theta_t\,(y^t)\, a_t\,(y^t, a_{-1}, s^t)\, \mathbf{\Lambda}\,(y_0, da_{-1}) = A_t(s^t) \equiv A_t$, where we omit the dependence on the history of aggregate states s^t for sake of simplicity.

equations (7)–(9);

2. *financial, labor, and goods markets clear at all dates: for any $t \geq 0$, equation (12) hold;*
3. *the government budget is balanced at all dates: equation (5) 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 (1), and post-tax definitions (4).*

4 The Ramsey problem

This paper aims at characterizing social welfare weights that allow the observed fiscal policy to be optimal. To do so, we need to compute for an arbitrary aggregate welfare function the fiscal policy that maximizes aggregate welfare. As mentioned in the introduction, various specifications of a Social Welfare Function (SWF) in heterogeneous-agent models are used in the literature. We first provide the form we use and then present the Ramsey program. The goal of the following construction is to build a SWF, where the planner puts some weights on the current productivity of agents, as in Heathcote & Tsujiyama (2021) in a static framework.

4.1 The Social Welfare Function.

In the sequential representation, the expected welfare of an agent having initial state (y_0, a_{-1}) can be expressed as follows:

$$V_0(y_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{y^t \in \mathcal{Y}^{t+1}|y_0} \theta_t(y^t \mid y_0) \left(u(c_t(y^t)) - v(l_t(y^t)) + u_G(G_t) \right),$$

where for the sake of simplicity, we omit the dependence on the initial wealth level a_{-1} and the aggregate shock s^t. We assume that the planner considers puts some weight $\omega(y_t)$ to the period utility function of agents having an idiosyncratic state y_t at period t. In this general formulation, the planner considers the expected intertemporal utility $V_0^P(y_0, a_{-1})$ as:

$$V_0^P(y_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{y^t \in \mathcal{Y}^{t+1}|y_0} \omega(y_t)\theta_t(y^t \mid y_0) \left(u(c_t(y^t)) - v(l_t(y^t)) + u_G(G_t) \right).$$

The SWF at date 0 then becomes:

$$W_0 = \sum_{y_0 \in \mathcal{Y}} \theta_0 (y_0) V_0^P (y_0),$$

which is the sum of period-0 intertemporal welfare of all agents. After simple algebra and using the law of large number, one finds that the planner maximizes at date 0 the following aggregate welfare:

$$W_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{y^t \in \mathcal{Y}^{t+1}} \theta_t(y^t)\omega(y_t) \left(u(c_t(y^t)) - v(l_t(y^t)) + u_G(G_t) \right).$$

Compared to the (unweighted) utilitarian Social Welfare Function, the only difference is the set of weights $\omega(y_t)$, that affect period utility function depending on the current idiosyncratic level. The number of Pareto-weights is thus finite and equal to the number of idiosyncratic states,

$Y = card\,(\mathcal{Y})$. Pareto weights are defined up to a positive scaling parameter. So, without loss of generality, we further impose the normalization constraint: $\sum_{y\in\mathcal{Y}} \omega\,(y) = 1$.

Following the notation of Remark 1, we will write compactly the SWF as

$$W_0 := \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \omega(y_t^i) \left(u(c_t^i) - v(l_t^i) + u_G(G_t)\right) \ell(di)\right]. \tag{14}$$

The objective of our paper is to estimate the Pareto weight of the SWF that enable the model to replicate: (i) the actual fiscal policies (in France or in the US); and (ii) some key empirical steady-state moments.

4.2 The Ramsey program.

In the Ramsey program, the planner aims at determining the fiscal policy corresponding to the competitive equilibrium – as specified in Definition 1 – that maximizes aggregate welfare according to the criterion of equation (14) – while satisfying the governmental budget constraint. Formally, the Ramsey program can be stated as follows.

$$\max_{\left(w_t, r_t, \tilde{w}_t, \tilde{r}_t, \tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t, G_t, K_t, L_t, (c_t^i, l_t^i, a_t^i, \nu_t^i)_i\right)_{t\geq 0}} W_0, \tag{15}$$

$$G_t + (1 + r_t)B_{t-1} + w_t \int_i (y_t^i l_t^i)^{1-\tau_t} \ell(di) + r_t K_{t-1} = \tau_t^c C_t + F(K_{t-1}, L_t, s_t) + B_t, \tag{16}$$

$$\text{for all } i \in \mathcal{I}\text{: } (1 + \tau_t^c)c_t^i + a_t^i = (1 + r_t)a_{t-1}^i + w_t(y_t^i l_t^i)^{1-\tau_t}, \tag{17}$$

$$a_t^i \geq -\bar{a},\ \ \nu_t^i(a_t^i + \bar{a}) = 0,\ \ \nu_t^i \geq 0, \tag{18}$$

$$u'(c_t^i) = \beta \mathbb{E}_t \left[(1 + r_{t+1}) \frac{1 + \tau_t^c}{1 + \tau_{t+1}^c} u'(c_{t+1}^i)\right] + \nu_t^i, \tag{19}$$

$$v'(l_t^i) = \frac{(1 - \tau_t)}{1 + \tau_t^c} w_t y_t^i (y_t^i l_t^i)^{-\tau_t} u'(c_t^i), \tag{20}$$

$$K_t + B_t = \int_i a_t^i \ell(di),\ \ L_t = \int_i y_t^i l_t^i \ell\,(di)\,, \tag{21}$$

and subject to several other constraints (which are not reported here for space constraints): the definition (1) of pre-tax wage and interest rates $\tilde{w}_t$ and $\tilde{r}_t$, the definition (4) 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 (17), individual credit constraint (and related constraints on ν_t^i) (18), Euler equations for consumption and labor (19) and (20), and market clearing conditions for financial and labor markets (21). Moreover, the fiscal policy selected by the Ramsey equilibrium should also fulfill the governmental budget constraint (16) – that is also a constraint.

We solve this program using a Lagrangian approach, that has been developed in LeGrand & Ragot (2022).[8] One advantage of this solution method is to simplify the derivation and the interpretation of the first-order conditions of the Ramsey program.

[8] LeGrand & Ragot (2022) show that this method is compatible with occasionally binding credit constraints. Açikgöz et al. (2018) propose another method to solve for policies with a utilitarian social welfare function.

4.3 Identification of Pareto weights.

The optimal Ramsey program follows the standard definition, which aims at finding the optimal fiscal policy while social preferences are exogenously given to. Our paper focuses here on an inverse optimal taxation problem. Observing actual market equilibrium and actual fiscal system, $(\tau_t^c, \tau_t^K, \tau_t, \kappa_t, B_t, G_t)_{t\geq 0}$, we aim at determining the *consistent social weights* that enable actual observations to be optimal allocations of the Ramsey program.

Definition 2 (Consistent SW)**.** *Consider a steady-state fiscal policy $(\tau^c, \tau^K, \tau, \kappa, B, G)$. Social Pareto weights implied by the mapping $\omega : \mathcal{Y} \to \mathbb{R}$ are said to be consistent when the fiscal policy $(\tau^c, \tau^K, \tau, \kappa, B, G)$ is the steady-state optimal fiscal policy of the Ramsey program with the planner's objective (14) defined using the weights implied by ω.*

The previous definition does not imply either existence or uniqueness of consistent social weights. Indeed, as can be expected, a large set of consistent social weights are compatible with a given fiscal policy that only includes those fiscal instruments. Social weights will thus be generically under-identified.

We use two distinct identification strategies. The first one consists in selecting among the set of consistent social weights the ones which are the closest – in the sense of the Euclidean norm – to the standard Utilitarian social welfare function. Indeed, at least since the work of Aiyagari (1995), the Utilitarian social welfare function is the standard benchmark in the heterogeneous agent literature – as we discussed in the literature review above. More formally, we define the *identified* consistent social weights, as the ones which minimize in each period the following distance:

$$(\omega(y))_{y\in\mathcal{Y}} = \arg\min_{\{\tilde{\omega}(y))\}} \sum_{y\in\mathcal{Y}} \theta(y)(\tilde{\omega}(y) - 1)^2, \tag{22}$$

subject to the constraint that the weights are consistent in the sense of Definition 2.

The second identification strategy consists in estimating a parametric functional form for $\omega(\cdot)$ – similarly to Heathcote & Tsujiyama (2021) – for as many degrees of freedom in the functional form as the number of fiscal policy constraints. More precisely, we retain the following function form:

$$\log \omega(y) := \omega_0 + \omega_1 \log(y) + \omega_2 (\log(y))^2,$$

where ω_0, ω_1, and ω_2 are three free parameters that are exactly identified with fiscal policy restrictions of Definition 2.

5 Characterizing the optimal fiscal policy

5.1 An redundancy result

Before solving for the Ramsey problem (15)–(21), we state an irrelevance result in our incomplete-market economy with aggregate shocks.

Proposition 1. *In the Ramsey problem (15)–(21), where the planner can levy linear taxes on consumption and capital, progressive tax on labor, and raise one-period public debt, the consumption tax is redundant with other fiscal instruments.*

Furthermore, the credit constraint $\overline{a}$ can be set to zero without loss of generality.

This proposition states that the optimal fiscal policy is not uniquely pinned down by the Ramsey program. More precisely, if we set a given exogenous path for consumption tax, the Ramsey problem will uniquely characterize the path of public debt, labor and capital taxes. This redundancy result requires the existence of only three other instruments: public debt, linear tax on capital and progressive tax on labor.

The other result of Proposition 1 is that the credit-constraint $\overline{a}$ can be set to 0 – or actually to any finite value – without loss of generality. The result already exists in endowment economies and only requires a rescaling of income streams.[9]

We provide the proof here as it allows us to simplify the Ramsey program, before deriving its first-order condition. We start with the following definition:

$$\tilde{a}_t^i := \frac{a_t^i}{1+\tau_t^c}, \tag{23}$$

$$W_t := \frac{w_t}{1+\tau_t^c}, \tag{24}$$

$$R_t := \frac{(1+r_t)(1+\tau_{t-1}^c)}{1+\tau_t^c}, \tag{25}$$

which represents the asset choices in (23), the wage rate in (24), and the interest rate in (25). With this notation, the agent's budget and credit constraints become:

$$c_t^i + \tilde{a}_t^i = W_t(y_t^i l_t^i)^{1-\tau_t} + R_t \tilde{a}_{t-1}^i, \tag{26}$$

$$\tilde{a}_t^i \geq -\frac{\overline{a}}{1+\tau_t^c} := -\tilde{\overline{a}}. \tag{27}$$

Since taxes and prices are considered as given by agents, we can equivalently state their optimization program using the notation (23)–(25) and the constraints (26) and (27) rather than the original notation and the constraints (8) and (9). This modifies Euler equations (10)–(11) as follows:

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

$$v'(l_t^i) = (1-\tau_t) W_t y_t^i (y_t^i l_t^i)^{-\tau_t} u'(c_t^i).$$

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

$$\tilde{B}_t := \frac{B_t}{(1+\tau_t^c)}, \tag{28}$$

$$\tilde{A}_t := \frac{A_t}{1+\tau_t^c}. \tag{29}$$

With this new definitions, the financial market equilibrium given by (21) holds since we have $\tilde{A}_t = \int_i \tilde{a}(i) l(di)$. The governmental budget constraint (16), using the resource constraint

[9] Aiyagari (1994) discusses the value of $\overline{a}$ in an economy without aggregate shocks and define this value to be the borrowing limit, while Shin (2006) provides the same discussion with aggregate shocks. By setting this value to zero we can guarantee that the consumption remains positive in all states.

(13), can be expressed as follows:[10]

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

We finally define:

$$\hat{B}_t := (1+\tau_t^c)\tilde{B}_t - \tau_t^c \tilde{A}_t, \tag{30}$$

and obtain for the governmental budget constraint:

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

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

The Ramsey program can now be written in post-tax as follows:

$$\max_{\left(W_t,R_t,\tilde{w}_t,\tilde{r}_t,\tau_t^c,\tau_t^K,\tau_t,\kappa_t,\hat{B}_t,G_t,K_t,L_t,(c_t^i,l_t^i,\tilde{a}_t^i,\nu_t^i)_i\right)_{t\geq 0}} W_0, \tag{31}$$

$$G_t + W_t \int_i (y_t^i l_t^i)^{1-\tau_t}\ell(di) + (R_t - 1)\tilde{A}_{t-1} + \hat{B}_{t-1} = F(K_{t-1}, L_t, s_t) + \hat{B}_t, \tag{32}$$

$$\text{for all } i \in \mathcal{I}:\ c_t^i + \tilde{a}_t^i = W_t(y_t^i l_t^i)^{1-\tau_t} + R_t\tilde{a}_{t-1}^i, \tag{33}$$

$$\tilde{a}_t^i \geq -\tilde{\bar{a}},\ \nu_t^i(\tilde{a}_t^i + \tilde{\bar{a}}) = 0,\ \nu_t^i \geq 0, \tag{34}$$

$$u'(c_t^i) = \beta\mathbb{E}_t\left[R_{t+1}u'(c_{t+1}^i)\right] + \nu_t^i, \tag{35}$$

$$v'(l_t^i) = (1-\tau_t)W_t y_t^i (y_t^i l_t^i)^{-\tau_t}u'(c_t^i), \tag{36}$$

$$K_t + \hat{B}_t = \tilde{A}_t = \int_i \tilde{a}_t^i \ell(di),\ L_t = \int_i y_t^i l_t^i \ell\,(di)\,. \tag{37}$$

5.2 Solving the Ramsey program

We now derive the first-order conditions of the Ramsey program (31)–(37). The detailed computation can be found in Appendix B.

The two first Lagrange multipliers we introduce are denoted by $\beta^t\lambda_{c,t}^i$ and $\beta^t\lambda_{l,t}^i$, and are associated to the date-t Euler equation of agent i for consumption and labor, respectively. To simplify the FOCs expression and ease their interpretation, we also introduce the so-called *social valuation of liquidity for agent* i and that we denote by ψ_t^i. It is formally defined as:

$$\psi_t^i := \omega_t^i u'(c_t^i) - (\lambda_{c,t}^i - R_t\lambda_{c,t-1}^i - (1-\tau_t)W_t y_t^i (y_t^i l_t^i)^{-\tau_t}\lambda_{l,t}^i)u''(c_t^i). \tag{38}$$

The quantity ψ_t^i can be seen as the marginal valuation of consumption from the planner's point of view. It includes the (private) marginal utility of consumption $u'(c_t^i)$ and also takes into consideration the saving incentives from periods $t-1$ to t and from periods t to $t+1$. An extra consumption unit makes the agent more willing to smooth out her consumption between periods

[10] A complete derivation of this result can be found in Appendix A.

t and $t+1$ and thus makes her Euler equation more "binding". This more "binding"constraint reduces the utility by the algebraic quantity $u''(c_t^i)\lambda_{c,t}^i$, where $\lambda_{c,t}^i$ is the Lagrange multiplier of the agent's Euler equation at date t. The extra consumption unit at t also makes the agent less willing to smooth her consumption between periods $t-1$ and t and therefore "relaxes" the constraint of date $t-1$ – which is reflected in the quantity $R_t u''(c_t^i)\lambda_{c,t-1}^i$. Finally, the last term in $\lambda_{l,t}^i u''(c_t^i)$ reflects how an extra consumption unit affects agent's labor supply incentives.

Similarly to ψ_t^i, we introduce the social marginal cost of labor $\psi_{l,t}^i$, which is defined as:

$$\psi_{l,t}^i := \omega_t^i v'(l_t^i) + \lambda_{l,t}^i v''(l_t^i).$$

It is the parallel of ψ_t^i for labor supply.

Finally, the last key Lagrange multiplier is denoted μ_t and is applied to the budget constraint of the government. Hence, for the government, the marginal cost at period t of transferring resources to households is μ_t. We also denote by $\hat{\psi}_t^i$ *the net social valuation of liquidity for agent* i, that is equal to the liquidity benefit ψ_t^i, diminished by the liquidity cost μ_t:

$$\hat{\psi}_t^i = \psi_t^i - \mu_t.$$

We now turn to the various FOCs. After proper substitution, the Ramsey program (31)–(37) implies six sets of FOCs, with respect to individual savings $\tilde{a}_t^i$, labor supply l_t^i, interest rate R_t, wage rate W_t, progressivity rate τ_t, and public debt $\hat{B}_t$, respectively.

The FOC with respect to individual savings ($\tilde{a}_t^i$) can be expressed as:

$$\hat{\psi}_t^i = \beta \mathbb{E}_t \left[R_{t+1} \hat{\psi}_{t+1}^i \right], \tag{39}$$

for agents who are not credit-constrained. For credit-constrained agents, we have $\lambda_{c,t}^i = 0$. Equation (39) is a Euler equation for social liquidity valuations, reflecting that the latter should be smoothed over time. This is a social version of the individual consumption Euler equation to ψ_t^i, which is itself the social version of the marginal utility of consumption.

The FOC with respect to labor (l_t^i) is:

$$\begin{aligned} \psi_{l,t}^i = {} & (1-\tau_t) W_t (y_t^i)^{1-\tau_t} (l_t^i)^{-\tau_t} \hat{\psi}_t^i \\ & + \mu_t F_{L,t} y_t^i - (1-\tau_t) W_t (y_t^i)^{1-\tau_t} (l_t^i)^{-\tau_t} \lambda_{l,t}^i \tau_t \frac{u'(c_t^i)}{l_t^i}, \end{aligned} \tag{40}$$

which is a generalized version of the labor Euler equation, where the ψs substitute for marginal utilities and where the general equilibrium effect of labor choices are accounted for.

The FOC with respect to the interest rate (R_t) can be written as:

$$\int_j \hat{\psi}_t^j \tilde{a}_{t-1}^j \ell(dj) = - \int_j \lambda_{c,t-1}^j u'(c_t^j) \ell(dj). \tag{41}$$

A marginal increase in the interest rate at t involves an increase in consumption for agent i proportional to her beginning of period asset holding $\tilde{a}_{t-1}^i$ and is thus valued by planner to $\hat{\psi}_t^i \tilde{a}_{t-1}^i$. However, this creates distortions on savings incentives from $t-1$ to date t by affecting agent's Euler equations, valued by the planner with $\lambda_{c,t-1}^i$. Hence, equation (41) sets equal the aggregate (on the whole population) on an increase in R_t to its aggregate distortions.

The FOC with respect to the wage rate (W_t) is:

$$\int_j \hat{\psi}_t^j (y_t^j l_t^j)^{1-\tau_t} \ell(dj) = -\int_j \lambda_{l,t}^j (y_t^j l_t^j)^{1-\tau_t} (1-\tau_t) u'(c_t^j)/l_t^j \ell(dj). \tag{42}$$

The interpretation is similar to equation (41). The benefit for agent i of an increase in W_t is proportional to her efficient labor supply $y_t^i l_t^i$ and is valued by the planner with $\hat{\psi}_t^i$. This implies the left hand-side of (42), while the right hand-side reflects the general equilibrium (through labor disincentives for instance) of raising W_t.

The FOC with respect to progressivity (τ_t) is:

$$\begin{aligned} 0 = & \int_j (y_t^j l_t^j)^{1-\tau_t} (\hat{\psi}_t^j + \lambda_{l,t}^j (1-\tau_t)(u'(c_t^j)/l_t^j)) \ln(y_t^j l_t^j) \ell(dj) \\ & + \int_j \lambda_{l,t}^j \left((y_t^j l_t^j)^{1-\tau_t}\right) (u'(c_t^j)/l_t^j) \ell(dj). \end{aligned} \tag{43}$$

The FOC with respect to the public debt ($\hat{B}_t$) is:

$$\mu_t = \beta \mathbb{E}_t \left[(1+\tilde{r}_{t+1})\mu_{t+1}\right]. \tag{44}$$

As FOC (39), this involves a Euler equation for the Lagrange multiplier μ_t on the governmental budget constraint. Note that the interest rate in (44) differs from the one in (39) and involves the pre-tax rate and not the post-tax one.[11]

Steady state. The steady-state characterization of LeGrand et al. (2021) for the fiscal policy still holds in this set-up. First, the steady-state pre-tax rate can still be deduced from the FOC (44) on public debt:

$$\beta(1+\tilde{r}) = 1, \tag{45}$$

which, given the CRS property of the production function, pins down the capital-to-labor ratio K/L. Second, the capital tax must satisfy:

$$\tau^K = \frac{\int_i \nu^i \ell(di)}{(1-\beta)\int_i u'(c^i)\ell(di)}, \tag{46}$$

where we recall that ν^i is the Lagrange multiplier on agent i's credit constraint. In other words, equation (46) states that the capital tax is proportional to the severity of the credit constraint.

In Appendix C we discuss the method we use to solve the Ramsey problem discussed in 5.1 as well as the technique we use to compute the social weights in (22).

[11]When the government spending affect the utility of agents, the marginal cost for the government to transfer resources to households at period t will be $\mu_t = \int_j \omega_t^j u'_G(G_t)\ell(dj)$.

6 Quantitative investigation

We first provide two calibrations to reproduce the tax system and the wealth distribution both in the US and France for the period 1995-2007. We then estimate Pareto weights for both economies and compare the implied Social Welfare Functions in both countries. We find the Pareto weights such that the US has the French tax system and compare how it evolves. Finally, we analyse the dynamics of the fiscal system of both countries and, particularly, the US business cycles properties when it has its SWFs altered.

6.1 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 $\beta = 0.992$ so as to match a realistic capital-to-output ratio. The period utility is specified such that $U(\cdot) = \frac{c^{1-\gamma}-1}{1-\gamma} + \frac{1}{\chi}\frac{l^{\frac{1}{\varphi}+1}}{\frac{1}{\varphi}+1}$ with $\gamma = 1.8$. Furthermore, the Frisch elasticity for labor is set to $\varphi = 0.5$, which is recommended by Chetty et al. (2011) for the intensive margin. We set the labor-scaling parameter to $\chi = 0.0477$, which implies normalizing the aggregate labor supply to 0.33.

Technology and TFP shock. 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 $\delta = 2.5\%$. The TFP process is a standard AR(1) process with $s_t = e^{z_t}$ and $z_t = \rho_z z_{t-1} + \varepsilon_t^z$, where $\varepsilon_t^z \sim_{\text{IID}} \mathcal{N}(0, \sigma_z^2)$. We ue standard values $\rho_z = 0.95$ and $\sigma_z = 0.31\%$ to obtain a deviation of the TFP shock around 1 % in a quarterly frequency.

Idiosyncratic labor risk. Various estimations of the idiosyncratic process can be found in the literature. The productivity follows an AR(1) process: $\log y_t = \rho_y \log y_t + \varepsilon_t^y$, with $\varepsilon_t^y \sim_{\text{IID}} \mathcal{N}(0, \sigma_y^2)$. The calibration features an autocorrelation $\rho_y = 0.99$ 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, which is the same as the one reported in Table 1. We obtain a Gini of wealth of 0.78. The model does a good job in matching the income and wealth distribution in the US.

Taxes and government budget constraint. Fiscal parameters are calibrated based on the computations by Trabandt & Uhlig (2011) reported in Table 1, with exception of the progressivity of the labor tax, which was computed using our own calculations based on the Luxembourg Income Study Database (LIS) and summarized 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.[12]

Finally we estimate the parameters κ such that it matches the public spending over GDP ratio. By doing this we use a value of $\kappa = 0.85$, which is close to the estimates in Ferriere & Navarro (2020). With this fiscal system, the model generates a public-debt-to-GDP ratio equal to 63%, which is the same value as in Table 1. The model also implies a public-spending-to-GDP ratio equal to 15%, the same as the one reported in Table 1.

In addition, the model does also good job in replicating the ratios of consumption over GDP and of investment over GDP. The model predicts a consumption-to-GDP ratio of 58%, which is 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%, which is close to the empirical value calculated to be 25%. Finally, regarding inequalities, the model generates a Gini index for post-tax income equal to 0.40, which is the same to its empirical counterpart provided in Table 1. The Gini index for wealth is found to be 0.78, which is very close to its empirical value of 0.77 in Table 1.

We gather the model implications in Table 4. These implications show that our tax system provides a good approximation of the redistributive effects of the actual tax system. This confirms the results of Heathcote et al. (2017) and Dyrda & Pedroni (2018).

6.2 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 in Section 6.1, even though our presentation is more streamlined. The calibration parameters can be found, as those for the US in Table 3.

[12]To estimate this parameter we use the dataset for the US in 2005 and we restrict our attention to the head of the households and their spouses with age between 25 and 60 and who were employed. We define *labor income* as being the sum of wage income, self-employment income, and private transfers. We then define a variable *net tax* as being the income tax and contributions minus public transfers and occupational pensions. The capital income is the sum of interest and dividends, rental income, and private pensions. We then apply the estimate of Trabandt & Uhlig (2011) of $\tau^K = 36\%$ in the capital income above mentioned and obtain the *capital income tax.* Finally, the *labor income tax* is simply the variable *net tax* minus *capital income tax.* We then proceed by defining the *disposable income* as the *labor income* minus the *labor income tax* and we regress the log of disposable income in the log of labor income. The results are presented in Table 2.

		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	·	$\gamma = 1.8$	·	$\gamma = 1.8$
φ	Frish 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%	Krueger et al. (2018)	2.5%	Own calculations, INSEE
Productivity parameters					
σ^y	std. err. productivity	0.10	Gini for income	0.06	Fonseca et al. (2020)
ρ^y	autocorr. productivity	0.99	Gini for income	0.99	Fonseca et al. (2020)

Table 3: Parameter values

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 0.29.

Technology and TFP shock. 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 $\delta = 2.5\%$.[13] The TFP process is a standard AR(1) process with $s_t = e^{z_t}$ and $z_t = \rho_z z_{t-1} + \varepsilon_t^z$, where $\varepsilon_t^z \sim_{\text{IID}} \mathcal{N}(0, \sigma_z^2)$. We ue standard values $\rho_z = 0.95$ and $\sigma_z = 0.31\%$ to obtain a deviation of the TFP shock around 1 % in a quarterly frequency.

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. (2020). With the process, the model is also able to replicate a realistic level of wealth inequalities. The Gini of income after taxes and transfers in the model is 0.28, which is the same as the value reported in Table 1, whereas the Gini of wealth is estimated to be 0.68.

Taxes and government budget constraint. We use the values summarized in Table 1 for the French taxes, with exception for the labor tax that in our model progressive. We consider a capital tax of $\tau^K = 35\%$, a parameter $\tau = 0.23$ (estimated in the same way as estimated for the US using the dataset of France for the year 2005), 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 match the empirical public-spending-to-GDP of 24 %. The public-debt-to-GDP ratio amounts to 60%, as the value of Table 1. Regarding private consumption and investment, the model generates an aggregate private consumption equal to 44% of GDP, which is close to the empirical counterpart of 45% estimated by Trabandt & Uhlig (2011) for the period 1995-2007, while the investment amount to 31% of GDP to be compared to 31% for its empirical counterpart. Finally,

[13] We are keeping the same values as in United States to emphasize more the differences in the SWFs due to differences in the fiscal system of both countries.

in terms of inequalities, the model implies a Gini coefficient for post-tax income equal to 0.28, the same as the empirical value reported in Table 1. The Gini index for wealth generated by the model amounts to 0.68, which is also the same as the empirical counterpart. These elements prove that the tax system made of linear tax for capital and progressive tax for labor is very relevant from an empirical point of view.

		US		France	
Parameter	Description	Model	Data	Model	Data
Public finance aspects					
B/Y	Public debt-to-GDP ratio	63%	63%	60%	60%
G/Y	Public spending-to-GDP ratio	15%	15%	25%	24%
	Total tax revenues	16%	26%	25%	40%
Aggregate quantities					
C/Y	Aggregate consumption (share of GDP)	58%	60%	44%	45%
I/Y	Aggregate investment (share of 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.

6.3 Estimation of Pareto weights

As explained above, the estimation procedure identifies Pareto weights such that the first-order conditions of the planner are satisfied (i.e., equations (39) to (44)) and which are the closest to the utilitarian Pareto weights.

The main issue to identify those weights arise because the Ramsey problem of Section 5.1 involves a joint distribution across wealth and Lagrange multipliers, which leads to a high-dimensional object with a high number of difficulties for the program resolution, especially in the presence of aggregate shocks. For instance, this joint distribution affects the planner's instruments in a non-obvious way, which makes the methods based on perturbation of a well-identified steady-state not usable for solving such problems (as Reiter 2009, Boppart et al. 2018, Bayer et al. 2019 or Auclert et al. 2019).

Due to it, in order to identify the Pareto weights we use the Lagrangian approach method developed in LeGrand & Ragot (2022). Basically this method allows us to compute the steady-state allocation and derive a finite number of equations that can simulate by perturbation the dynamics of the Ramsey program for small aggregate shocks. The idea is to build an aggregation of the Bewley model (thus for a given policy and no aggregate shock) in which agents with the same history over last N (where N is a fixed horizon) periods are aggregated into a unique "agent". This method implies that the "aggregate" agent is endowed with the average wealth and average allocation of all individuals with this N-period history.

Let $N \geq 0$ be a truncation length. The key step of the aggregation consists in assigning to all agents sharing the same idiosyncratic history over the last $N \geq 0$ periods the same wealth and the same allocation. Such a N-period history will be said to be a truncated history and for a history $y^t = \{y_0, \dots, y_{t-N}, y_{t-N+1}, \dots, y_{t-1}, y_t\}$, this corresponds to the N-length vector denoted $y^N = \{y^N_{t-N+1}, \dots, y^N_{t-1}, y^N_t\}$. To sum up we can represent the truncated history of an agent i whose idiosyncratic history is y^t as:

$$y^t = \{\underbrace{y_0, \dots, y_{t-N-2}, y_{t-N-1}, y_{t-N}}_{\xi_{y^N}}, \underbrace{y^N_{t-N+1}, \dots, y^N_{t-1}, y^N_t}_{=y^N}\},$$

where the parameter ξ_{y^N} captures the residual heterogeneity for the truncated history y^N and is built such that the truncated model will be an exact aggregation of the underlying Bewley model in the absence of aggregate shocks. In Appendix C we present with more details the truncation method, the Ramsey program of Section 5.1 using the truncated model, as well as the FOC for the planner in this case. Finally we write the FOC at the steady state and using simple matrix algebra we pin down the weights either through the specification in equation (22) or using a parametric functional form for $\omega(.)$, similar to Heathcote & Tsujiyama (2021). In both cases we show that the Pareto weights satisfy the FOC of the planner.

For the exercises below we use a truncation length of $N = 5$, although the main characteristic of the results does not change when we consider a longer truncation length (see Appendix E). We select 10 idiosyncratic productivity levels, which implies $10^5 = 100000$ different truncated histories. The Pareto weights are estimated such that histories with the same productivity level in the beginning of the truncation will be assigned the same weight (i.e., if $y^N_t = \tilde{y}^N_t$ such that $y^N_t \in y^N$ and $\tilde{y}^N_t \in \tilde{y}^N$ with $y^N \neq \tilde{y}^N$ then $\omega(y^N) = \omega(\tilde{y}^N)$). In the end this means we will have a set of 10 Pareto weights, one for each possible value that the idiosyncratic variable can assume.[14]

We plot in Figure 2 the Pareto weights as a function of the productivity level. The weight for a given productivity level y_k is computed using the explanation above, while accounting for the population distribution.

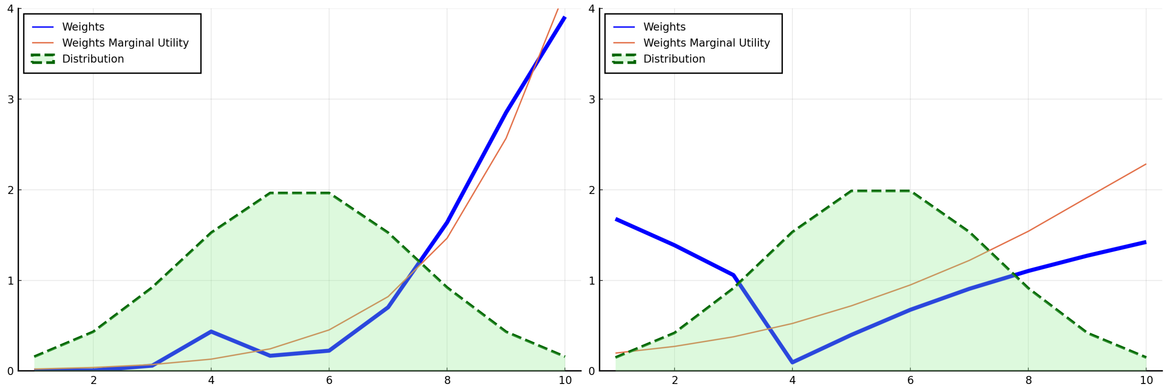
We can also consider an approach where we compute a Pareto weight for each history $k = 1, \dots, 100000$.[15] After calculate those weights, we can proceed in a similar fashion and get the productivity at the beginning of the truncation for each one of the histories, which will give us a pair denoted by $(y_k, \tilde{a}_k)$ where $\tilde{a}_k \in [-\tilde{\bar{a}}, \infty)$ is the wealth for the history k and $y_k \in \{y_1, \dots, y_{10}\}$ is the productivity level in the beginning of the truncation for each possible history k. We can then plot as a surface the mapping between the productivity-wealth pair $(y_k, \tilde{a}_k)$ and the corresponding Pareto weight. We represent the pareto weights using this strategy in Appendix D.[16]

Figure 2 plots the Pareto weights along the productivity dimension for the agents. We

[14] This result is stated in Appendix C, equation (C.55).

[15] This result is stated in Appendix C, equation (C.56). It is also possible for the planner to have Social Pareto Weights according to the initial distribution $(y_0, \tilde{a}_{-1})$, however, this is not consistent since agents with low values of productivity and wealth can become more productive and richer due to a sequence of good shocks, which makes the initial distribution to become less relevant.

[16] Figure 10 in Appendix D shows the Pareto weights for the US and France as a function of histories (i.e., such that for each $k = 1, \dots, 100000$ we have a Pareto weight) together with the Pareto weights such that histories with the same productivity in the beginning of the truncation are given the same weight (i.e., if $y^N_t = \tilde{y}^N_t$ such that $y^N_t \in y^N$ and $\tilde{y}^N_t \in \tilde{y}^N$ with $y^N \neq \tilde{y}^N$ then $\omega(y^N) = \omega(\tilde{y}^N)$). In this last case we will be let with 10 possible Pareto weights. The histories are organized in increasing order of the productivity level in the beginning of the truncation, which means this last case represents the weights in Figure 2. One can notice that although there are differences between those two ways to compute the weights, they are globally very similar.



(a) United States (b) France

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

can see that for the US the Pareto weights are increasing with the productivity level, whereas for France they are U-shaped, with higher weight for low productivity agents than for those at the top of productivity distribution. We also plot the weights considering the inverse of the marginal utility of those agents. Notice that higher productivity agents have lower marginal consumption and, hence, a higher weight.

In the US, the agents with the highest weight in the population are the high productivity agents, and the gap in weights between those agents and those with the lowest level of productivity amounts to around 3.9. In France, low-productivity agents have a higher weight than those with high productivity, although the gap between these two categories is not as sizeable as in the US (e.g., for France this value amounts to 0.25). Table 5 contains some summary statistics for the Pareto weights computed for France and the US.

	US	**France**
Mean	1.00	1.00
St. deviation	1.37	0.49
Min.	0.006	0.095
Max.	3.91	1.68
Bottom 10 %	0.006	0.37
Median	0.33	1.08
Top 10%	2.96	1.45

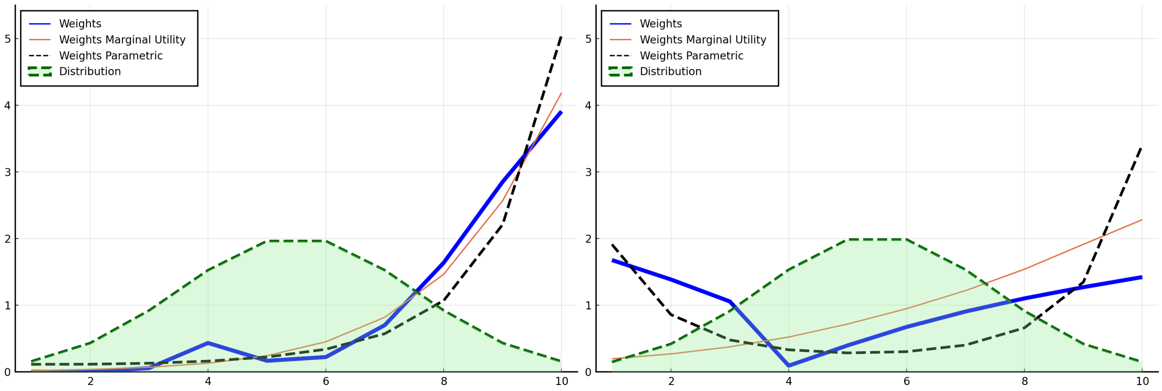
Table 5: Summary statistics for the Pareto Weights of the US and France.

In line with the literature we also compute a parametric version of the Pareto Weights. Those weights are obtained such that the first order conditions of the planner are still satisfied. In order to calculate those weights we consider the following functional form:

$$\log \omega(y) := \omega_0 + \omega_1 \log(y) + \omega_2 (\log(y))^2,$$

where y represents the productivity states. More details about this calculation can be found in the Appendix C, but basically the idea is that the FOC of the planner will give us three sets of conditions to be met and we use those three conditions to exactly identify the three parameters above.[17]

[17] We use the restrictions in the problem stated in Appendix C, equation (C.55) to exactly identify the parameters



(a) United States (b) France

Figure 3: Parametric Pareto weights as a function of productivity for the US and France.

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 3 we plot the parametric weights together with the weights we had plotted previously.

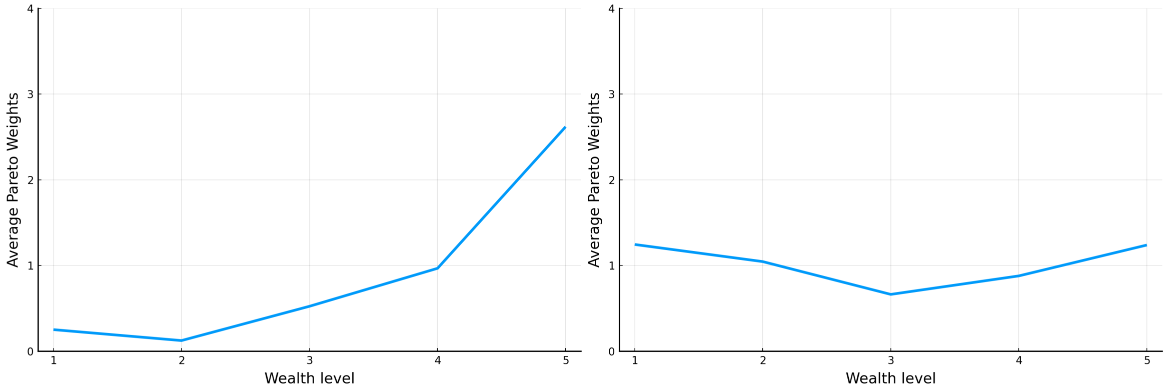
Notice that the parametric weights has the same shape we had obtained through direct estimation of the weights, however, for the US the difference in weights between the high productivity and low productivity agent become more pronounced. For France, we can see that the weights for the high productivity agents are now higher than the weights for the low productivity ones, although we still have the U-shape format and similar weights for the low-productivity agents.

Finally, we report in Figure 4 the average Pareto weights along the wealth dimension. We compute these by averaging the values of the Pareto weights for the individuals who are in the bottom 10 % (level 1), between the bottom 10 % and 20 % (level 2), 20 % and 50 % (level 3), 50 % and 90 % (level 4), and for those who belongs to the top 10 % of the wealth distribution (level 5).

We can observe that the French shape is similar to the one we had obtained when we calculate the Pareto weights along the productivity dimension. It can also be observed, as for income, that the gap in Pareto weights between highest and lowest weights is higher in the US than in France. For the US although the weights are still increasing, we can observe a higher value of weight for the agents at the bottom of wealth distribution, which means this dimension is more important in the US than in France. However, in the Appendix E when we plot the results by higher values for N, we can see that although US still has an increasing weight in the wealth dimension, the same pattern does not continue to appear in the French case. In other words, for France it seems the wealth dimensions is way less relevant to explain the weights the social planner gives for the individuals.

As individual welfare is higher for low income agents and the wealth dimension seems to have less power to explain the weights in France than in the US, the French shape is consistent

ω_0, ω_1, and ω_2.



(a) United States (b) France

Figure 4: Average Pareto weights as a function of wealth for the US and France.

with a inequality or inequity aversion. The US shape is consistent with a strong redistribution component for low-wealth agents, but other favors high-income and high-wealth agents, at the expense of middle-class. As a final remark, our methodology is agnostic about the motivation of the planner behind these weights. They can result of various considerations regarding justice or moral choices, such as compassion, inequality aversion, or meritocracy, among others. We see this flexibility as an advantage of our methodology.

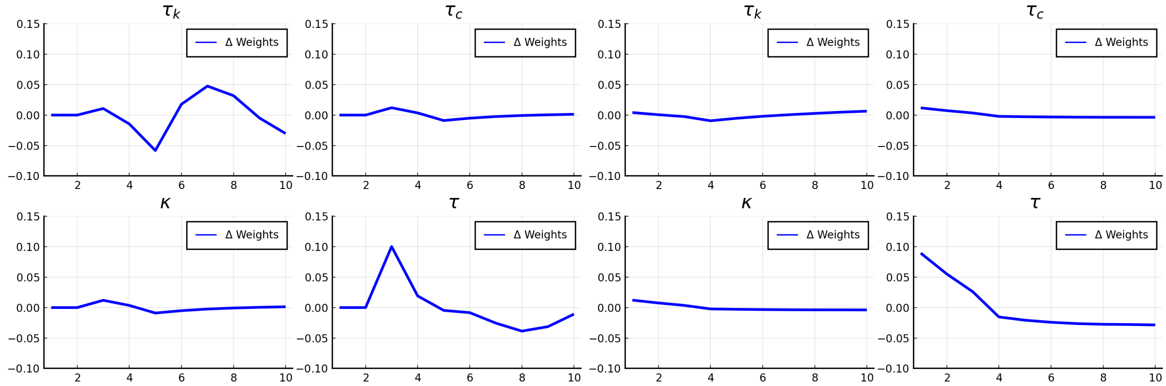
In Figure 5 we test the robustness of our estimates through a simple exercise. We increase the government spending to output ratio of United States by 10% and of France in 1%. We then conduct different experiments financing this increase by each one of the instruments (i.e., τ_k, τ_c, κ, and τ) such that the budget constraint of the state is still satisfied. For each one of the instruments we plot the difference in the weights using our estimation strategy discussed previously. Table 6 summarizes the changes in the instruments such that the budget of the state is satisfied after an increase in G/Y. Notice that this increase in G/Y is financed either by an increase in the tax on capital (τ_k), tax on consumption (τ_c), average tax on labor ($1-\kappa$), or progressivity (τ). It is worth noting that finance the increase in G by either τ_c or κ has the same effect in terms of general equilibrium, which implies the effect in terms of difference in weights is the same for both experiments. This result is linked to the redundancy result we discussed in Section 5.1.[18]

Finally, one can realize that the instrument that leads to the highest change in the weights is the progressivity of the labor tax, which corroborates to the idea that by increasing progressivity we benefit agents with low productivity shocks at the expense of those at the top, although even in this case the Pareto weights keep the same shape as the ones we show in Figure 2. In Appendix F, Figures 21 and 22 illustrate that the shape of the Pareto weights does not change for each experiment under consideration, being the effect almost negligible for the cases where we finance the increase in G/Y by τ_k, τ_c, and κ. This is because the effects in terms of general equilibrium are not drastic and as a result the planner does not alter the weights in a substantial way.

[18]In Appendix F we present the same difference in weights for the parametric estimation, as well as the change in terms of the utility, labor, and capital income for each experiment under consideration. This result is stated in Figure 19 for US and Figure 20 for France.

	US		France	
	Steady state	Increase in G/Y	Steady state	Increase in G/Y
τ_k	0.36	0.387	0.35	0.361
τ_c	0.05	0.076	0.18	0.19
κ	0.85	0.83	0.728	0.72
τ	0.16	0.22	0.23	0.25

Table 6: Changes in the fiscal instruments after an increase in G/Y for United States and France.



(a) United States (b) France

Figure 5: Change in weights by increasing G/Y for United States and France.

6.4 A world where United States has the French tax system

We use the previous methodology to investigate how the Pareto weights of the US behave when the US economy adopts a fiscal system similar to the French one. The motivation for this exercise is twofold. In one direction we wish to evaluate whether the fiscal system in France, with a higher progressivity in labor tax can explain the higher weights that France gives for low productivity agents, since the two fiscal systems seems to have different goals, with the French being consistent with a inequality or inequity aversion and the one of the US consistent with a more strong redistribution component for low-productivity agents when they are not wealthy.

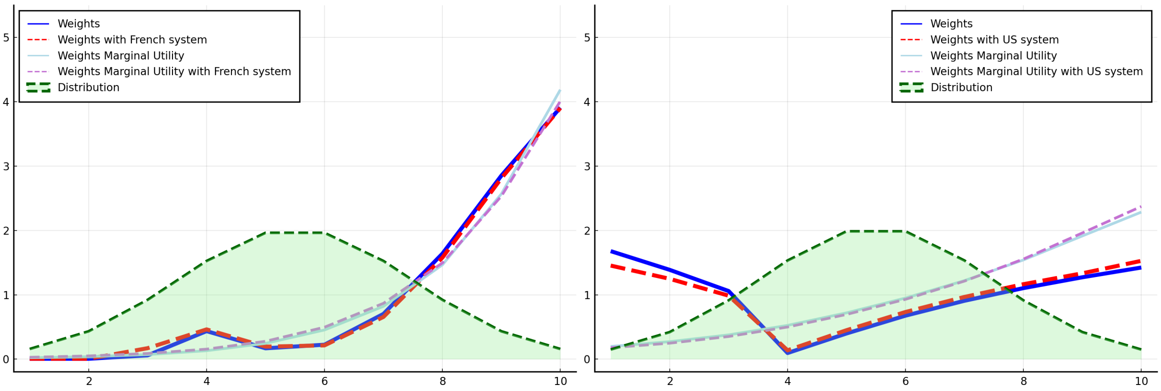
In order to obtain higher Pareto weights for low productivity agents we expect that higher progressivity can be helpful, while at the same time it helps to reduce the weights for the high productivity agents.

Secondly, we wish to analyse whether the business cycle properties are altered if the SWF of the US is different, and in which dimension those properties are altered. The goal here is to evaluate from the point of view of the Social Planner whether he is willing to accept the costs in terms of the business cycle by moving to a different social norm.

The first experiment is conducted as follows. Once the capital-to-output ratio is set to the value in the steady state, we iterate in the value for κ such that the value of government spending to output ratio is kept the same. By running the experiment in this way not only the model parameters are unchanged but also the main macro ratios. In this exercise the only vector we are changing is the vector that represents the fiscal system $(\tau_K, \tau_c, \tau, \kappa, B)$.[19]

In panel (a) of Figure 6 we plot the the Pareto weights as a function of productivity for

[19] Table 11 in Appendix G shows the tax system for United States and France that we obtain as a result of this experiment.



(a) United States (b) France

Figure 6: Change in the Pareto weights for United States and France by adopting a different tax system.

the US with the French tax system together with the Pareto weights in the steady state. In panel (b) we plot the case for France when it adopts the US tax system and the steady state one. It is straightforward to notice that when the US adopts the French system the weights for agents at the bottom increase at the expenses of those with higher productivity. The opposite is true when France adopts the US tax system. In this case, the weights for the low productivity agents is reduced and increased for those with high productivity.

In Appendix G, Figure 23 we show how the Pareto weights, utility, labor, and capital income change when we alter the fiscal system in the US for the French one and vice-versa. One can notice that the adoption of the French tax system for the US reduces the labor income for high productivity agents. This comes from the fact that the adoption of a more progressive tax system reduces the incentives for high productivity agents to supply labor. Moreover, since the income of those agents is now being more heavily taxed this leads to a decrease in the utility and in a reduction of the weights that the social planner gives for them. On the other hand, one can perceive that the more progressive tax system increase the consumption levels of low productivity agents and increase their utility, increasing the weights the social planner gives for them.

This experiment shows first that one can rationalize consistent changes in the tax system by a change in the Pareto weights, in the sense, that even for the US economy the adoption of those weights highlighted in dotted red in Figure 6 leads to a tax system equal to the French one. Consistency requires that the before tax interest rate is kept at its optimal value, which is the inverse of the discount factor, as well as the main parameters in the model. We could see that in order for the US to increase the weights for low productivity agents and decrease the weights for high productivity agents, one way for this to happen is through the adoption of a more progressive labor tax, since the relatively low labor tax in US compared to France favors the high-income/high-productivity agents.

Admittedly, as we compare steady-states, these are long-run effects, which do not consider the additional welfare effects of the transition from the benchmark fiscal system to the new one.

6.5 A world where United States has the Pareto weights of France

Now we investigate the fiscal system in the United States such that the United States has the Pareto weights as closest as possible to the French ones. The idea here is to find the fiscal system such that the Euclidean distance between the Pareto weights is the smallest possible. As the previous exercise we wish first to evaluate how the United States should behave in terms of the fiscal system such that it mimics in part the French system, since the two SWFs seem to have different goals. Secondly, we aim to analyse whether the business cycle properties are altered if the SWF of the United States is different and in which dimension those properties are altered. The goal here is to evaluate from the point of view of the Social Planner whether he is willing to accept the costs in terms of the business cycle by moving to a different social norm.

We conduct the experiment as follows. Once the capital-to-output ratio has been set to the value in the steady state, for every combination of the tax on capital and progressivity of the labor tax we iterate on κ such that the value of government spending to output ratio is kept the same. By doing this we are able to match the main macro ratios. More precisely, the model parameters are unchanged as well as the main macro ratios. In every case the only vector we are changing is the vector that represents the fiscal system $(\tau_K, \tau_c, \tau, \kappa, B)$. For each combination we calculate the Pareto weights and by using interpolation on the productivity states of the target economy (i.e., in this case France), we can obtain the Pareto weights on the union of the grids that represents the whole set of the productivity states. Once we obtain these Pareto weights we can calculate the distance between the vectors using standard measures as the Euclidean distance. We then consider, as in our benchmark case, that this new fiscal system that represents the minimum distance between the Pareto weights results from optimal planner's decisions.

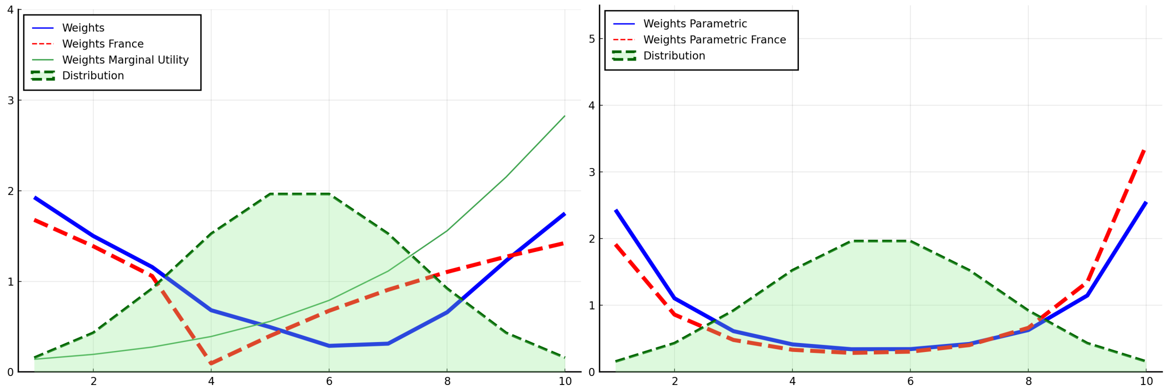
In this experiment we are considering the distance between the Pareto weights on the union of the grids in the productivity states, since as we argued above the Pareto weights is much less sensitive in the wealth dimension. Another experiment we also did was to calculate the distance between the Pareto weights for the possible histories $k = 1, \dots, 100000$ but the result we obtained below is kept the same.

For the United States, the new values of the capital and progressivity for the labor tax amount to 9 % and 57 % respectively, compared to 36 % and 16 % in the initial economy. Table 7 presents the values of the new fiscal system and compare it to the USA benchmark economy and the French benchmark economy.

US						
	Public debt (%GDP)	τ_k (%)	τ (%)	κ (%)	Gini a.t.	Gini wealth
Benchmark economy USA	63	36	16	85	40	78
Benchmark economy France	60	35	23	72.8	28	68
USA with the French PWs	299	9	57	71	27	63

Table 7: Comparison between the benchmark economies and the USA economy with the French Pareto weights.

The column Gini a.t. represents the Gini after taxes and transfers, and the last one represents the Gini for wealth. Observe that the distribution of income and wealth for the



(a) Pareto weights (b) Pareto weights parametric

Figure 7: Pareto weights for the United States with the French Pareto weights and France.

United States is now closer to the French one.

In this new scenario we have that the debt-to-GDP ratio increases to 299 %. This last result was expected, since for the distance to be minimized we need to decrease the weights on high income agents for the United States. The idea here is that the increase in public debt decreases relatively the weights in high income agents, as the state has to increase the overall labor tax to balance the budget. As expected the overall labor tax increased. The intuition of this result is that previously the labor tax for the United States was low, which favors the high-productivity/high-income agents

In this new scenario the capital goes down from 36 % to 9 % since savings need to absorb the additional supply of public debt. Notice that for the weights in the low-productivity/ low income agents to increase we need to enlarge the progressivity. The result corroborates with this intuition, with the progressivity being enlarged from 16 to 57 % . We recall that the other policy parameters, as well as the main macro ratios (e.g., consumption-to-GDP, government spending-to-GDP, investment-to-GDP) are the same as in the benchmark economy.[20]

Figure 7 plots the Pareto weights as a function of productivity for the US with the new fiscal system and France, where in panel (a) we have the Pareto weights considering the technique we discussed previously and in panel (b) we show the parametric version.

This experiment shows first that one can rationalize consistent changes in the tax system by a change in the Pareto weights. Consistency requires that the before tax interest rate is kept at its optimal value, which is the inverse of the discount factor, as well as the main parameters in the model. We could see that in order for the United States to mimic the Pareto weights for France we need an increase in the public debt, which in turn requires a decrease in capital taxes to generate additional savings and an increase in the labor tax to compensate for the loss in tax returns. Moreover, the increase in the labor tax was expected since the relatively low labor tax of the United States favors the high-income/high-productivity agents. This decreases the estimated weights for high-labor-income agents. Admittedly, as we compare steady-states, these are long-run effects, which do not consider the additional welfare effects of the transition from the benchmark fiscal system to the new one.

In the next sessions we analyse the dynamics of the economy and particularly how the

[20] In the Appendix H we have the evolution of the main variables in the model along the dimension (τ_K, τ).

US dynamics change when the SWFs is altered.

6.6 Dynamic of the fiscal system

We now present the dynamics of the fiscal system for both the US and France after a technology shock. To ease the comparison and be able to see the differences between the business cycle in both countries we keep the shock to be the same for both economies. From Proposition 1 we know that the tax on consumption does not play any role in this case. Due to it we do not show its dynamics, since it is assumed to be constant. Figure 8 compares the Impulse Response Functions (IRFs) for the US, France, for the case the US has the French tax system, and for the case France has the US tax system.

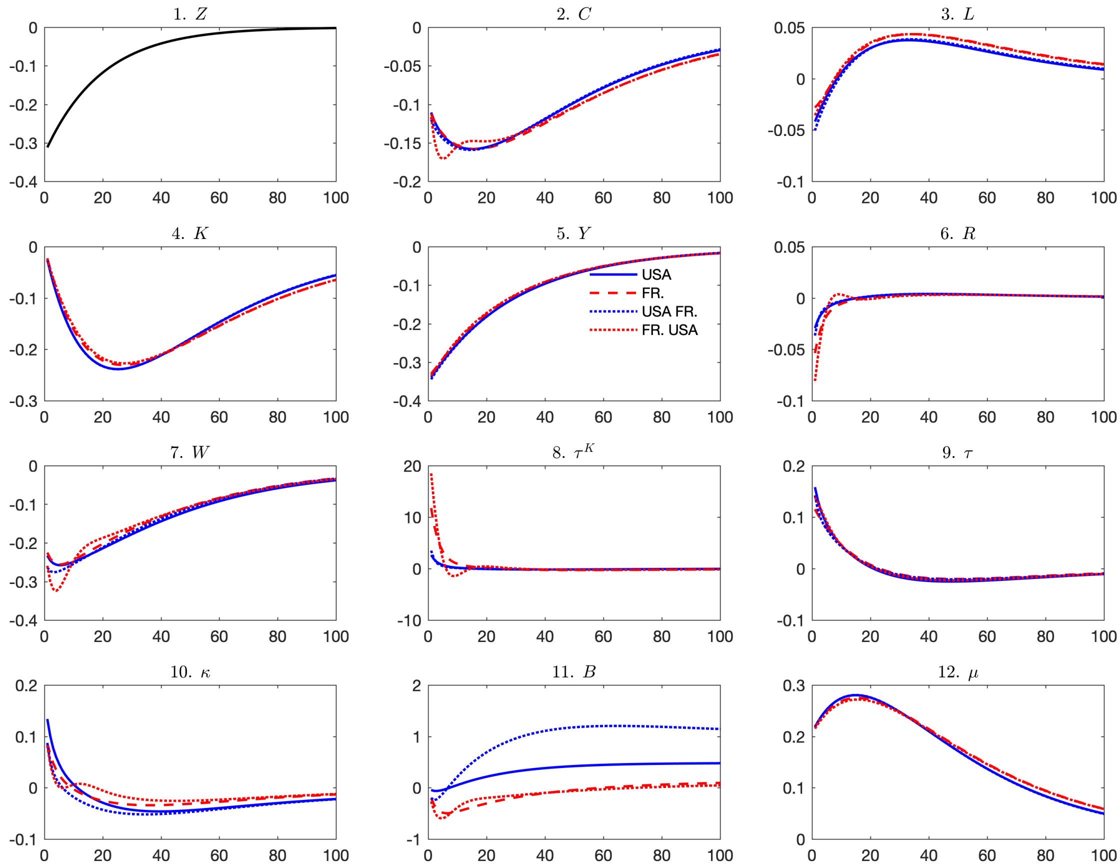


Figure 8: Comparison between the US economy (blue solid line), the French economy (red dotted line), the US economy with the French tax system (blue dashed line), and the French economy wit the US tax system (red dashed line).

Notice that each panel reports the proportional change for the variable under consideration, in percentage points. For instance, Panel 1 reports a persistent fall in TFP for 100 periods, after a fall of 0.31 % on impact, except for the tax rates, for which we compute the absolute variation.

We can see for the dynamics of the aggregate quantities (Consumption, Panel 2; Capital, Panel 4; Output, Panel 5) that both economies exhibit a similar behavior, with the French economy taking longer to reach the steady state and having a higher negative impact, especially in consumption. In overall, the U.S. economy is more flexible and the recover is faster than in France. When France has the US tax system we can observe the impact on consumption is higher in the first periods, but after a while it assumes a path similar to the benchmark case for France. For all the aggregate variables considered here, the path is similar even when we change the Pareto weights.

For the prices we have similar behavior for both economies, with the value of the interest rate R decreasing, which can be explained by the increase in the tax on capital. One can notice that the impact in R is stronger in France, since its capital tax is much more volatile than in the US. The reasoning is as follows, the negative shock causes the government revenue to be reduced, then in order to compensate for it the planner needs to accumulate assets to pay for the public spending. Given the government spending in France is higher than in the US, the capital tax in the former needs to increase more than in the US.

For the US we have a more moderate increase in the capital tax. Notice that in both economies there is an increase in the average tax on labor to compensate the negative impact on public finance caused by the negative shock.

By analysing the business cycle variables of the US when it has the French tax system we can notice that in terms of the dynamics of the economy, the business cycle variables of the US is almost unchanged for the aggregate quantities (Consumption, Capital, and GDP).

The variables that change more are the are the interest rate, wage, tax on capital, progressivity, average tax on labor, and public debt. With exception for the public debt all the other variables have a similar behavior.

This result confirms in some sense, the intuition and the idea that the SWFs can be linked to the fiscal system of the country and the cost in terms of adopting a new social norm comes to the expense in accepting new dynamics for the fiscal variables.

Table 8 summarizes the fist and second moments of the main variables in the US benchmark economy, French benchmark economy, the US with the French tax system, and for France with the US tax system. For each variable we report the steady state value, reported as "Mean" and the normalized standard deviation, which is equal to the standard deviation divided by the mean, referred here as "Std". In the taxes we report simply the standard deviation. Analysing the results reported in the Table 8 we can see that the main changes that occur in United States when it has the French tax system is related to the volatility of the tax on capital and the public debt, which now is higher than in the Benchmark economy. For France, we can notice that when it adopts the US tax system the tax on capital becomes much more volatile and the public debt less volatile. For US the public debt is counter-cyclical and volatile, whereas for France is cyclical and less volatile.

		USA	France	USA w/ Fr.	Fr. w/ USA
C	Mean	0.739	0.538	0.727	0.549
	Std	0.010	0.011	0.010	0.011
L	Mean	0.332	0.295	0.326	0.301
	Std	0.003	0.003	0.003	0.003
K	Mean	13.655	15.068	13.428	15.365
	Std	0.016	0.016	0.016	0.016
Y	Mean	1.264	1.215	1.243	1.239
	Std	0.013	0.013	0.013	0.013
τ^K	Mean	0.360	0.350	0.350	0.360
	Std	0.039	0.174	0.044	0.249
τ	Mean	0.160	0.230	0.230	0.160
	Std	0.004	0.003	0.003	0.003
κ	Mean	0.848	0.728	0.974	0.648
	Std	0.007	0.003	0.006	0.002
B	Mean	3.464	2.945	1.256	4.533
	Std	0.261	0.028	0.548	0.022
		Correlations			
corr(τ^K,Y)		-0.424	-0.570	-0.367	-0.402
corr(τ,Y)		-0.448	-0.545	-0.512	-0.516
corr(κ,Y)		0.129	0.217	0.328	0.127
corr(B,Y)		-0.142	0.621	-0.129	0.772
corr(C,Y)		0.909	0.895	0.914	0.913
corr(Y,Y_{-1})		0.967	0.967	0.967	0.965
corr(B,B_{-1})		0.999	0.995	0.999	0.986

Table 8: First and second moments for key variables in the cases considered - Benchmark US, Benchmark France, US with the French tax system, and France with US tax system.

7 Conclusion

We derive a methodology to identify the Social Welfare Function (SWF) of a government, which is compatible with the empirical wealth and income distributions given the actual tax structure. By using this methodology we calculate the Pareto weights for France and the US. To calculate the Pareto weights we selected among the large set of possible SWFs the closest one to the Utilitarian SWF, which attributes the same weights to all agents. By using four fiscal instruments: consumption, capital and progressive labor taxes, and public debt we estimated the SWFs of the two countries and showed that they differ from each other. The SWF for France gives a higher weight for low productivity agents and it is less heterogeneous than those of the US, while the US SWF has an increasing shape in income with more weight to high income agents. Our methodology allows computing the optimal allocation with aggregate shocks. We assess the change in business cycle properties if the SWF changes and we show that the main changes occurs in the fiscal instruments, being the dynamics of the aggregate variables almost unchanged with the change in the SWFs.

References

Açikgöz, O. (2016), On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production, Working Paper, Yeshiva University.

Açikgöz, O., Hagedorn, M., Holter, H. & Wang, Y. (2018), 'The Optimum Quantity of Capital and Debt', *Working Paper* .

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

Aiyagari, S. R. (1995), 'Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting', *Journal of Political Economy* **103**(6), 1158–1175.

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

Auclert, A., Bardóczy, B. & Rognlie, M. (2021), 'MPCs, MPEs, and Multipliers: A Trilemma for New Keynesian Models', *Review of Economics and Statistics* **Forthcoming**.

Auclert, A., Bardóczy, B., Rognlie, M. & Straub, L. (2019), 'Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models'.

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

Bayer, C., Luetticke, R., Pham-Dao, L. & Tjaden, V. (2019), 'Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk', *Econometrica* **87**, 255–290.

Benabou, R. (2002), 'Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?', *Econometrica* **70**(2), 481–517.

Bhandari, A., Evans, D., Golosov, M. & Sargent, T. J. (2020), 'Inequality, Business Cycles, and Monetary-Fiscal Policy'.

Boppart, T., Krusell, P. & Mitman, K. (2018), 'Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative', *Journal of Economic Dynamics and Control* **89**, 68–92.

Bourguignon, F. & Amadeo, S. (2015), 'Tax-Benefit Revealed Social Preferences', *Journal of Economic Inequality* **1**(10), 75–108.

Chang, B.-H., Chang, Y. & Kim, S.-B. (2018), 'Pareto Weights in Practice: A Quantitative Analysis Across 32 OECD Countries', *Review of Economic Dynamics* **28**, 181–204.

Cheridito, P. & Sagredo, J. (2016), 'Existence of Sequential Competitive Equilibrium in Krusell-Smith Type Economies'.

Chetty, R., Guren, A., Manoli, D. & Weber, A. (2011), 'Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins', *American Economic Review* **101**(3), 471–475.

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

Dyrda, S. & Pedroni, M. (2018), 'Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks'.

Ferriere, A. & Navarro, G. (2020), The Heterogeneous Effects of Government Spending: It's All About Taxes, Working Paper, Paris School of Economics.

Fleurbaey, M. & Maniquet, F. (2018), 'Optimal Taxation Theory and Principles of Fairness', *Journal of Economic Literature* **56**(3).

Fonseca, R., Langot, F., Michaud, P.-C. & Sopraseuth, T. (2020), Understanding Cross-Country Differences in Health Status and Expenditures, NBER Working Paper 26876, National Bureau of Economic Research.

Gornemann, N., Kuester, K. & Nakajima, M. (2016), 'Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy'.

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

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

Heathcote, J. & Perri, F. (2017), 'Wealth and Volatility', *Review of Economic Studies* **forthcoming**.

Heathcote, J., Storesletten, K. & Violante, G. L. (2017), 'Optimal Tax Progressivity: An Analytical Framework', *Quarterly Journal of Economics* **132**(4), 1693–1754.

Heathcote, J. & Tsujiyama, H. (2021), 'Optimal Income Taxation: Mirrlees Meets Ramsey', *Journal of Political Economy* **129**(11), 3141–3184.

Kaplan, G. & Violante, G. L. (2014), 'A Model of the Consumption Response to Fiscal Stimulus Payments', *Econometrica* **82**(4), 1199–1239.

Krueger, D., Mitman, K. & Perri, F. (2018), 'On the Distribution of the Welfare Losses of Large Recessions', pp. 143–184.

LeGrand, F., Martin-Baillon, A. & Ragot, X. (2021), Should Monetary Policy Care About Redistribution? Optimal Fiscal and Monetary Policy with Heterogeneous Agents, Working Paper, SciencesPo.

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

Marcet, A. & Marimon, R. (2019), 'Recursive Contracts', *Econometrica* **87**(5), 1589–1631.

Mendoza, E., Razin, A. & Tesar, L. (1994), 'Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption', *Journal of Monetary Economics* **34**(3), 297 – 323.

Miao, J. (2006), 'Competitive Equilibria of Economies with a Continuum of Consumers and Aggregate Shocks', *Journal of Economic Theory* **128**(1), 274–298.

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

Saez, E. & Stantcheva, S. (2016), 'Generalized Social Marginal Welfare Weights for Optimal Tax Theory', *American Economic Review* **106**(1).

Shin, Y. (2006), Ramsey Meets Bewley: Optimal Government Financing with Incomplete Markets, Working Paper, University of Wisconsin.

Tauchen, G. (1986), ‘Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions’, *Economics letters* **20**(2), 177–181.

Trabandt, M. & Uhlig, H. (2011), ‘The Laffer Curve Revisited’, *Journal of Monetary Economics* **4**(58), 305–327.

Werning, I. (2015), Incomplete Markets and Aggregate Demand, NBER Working Paper 21448, National Bureau of Economic Research.

Appendix

A Transformation of the governmental budget constraint

Using the government budget constraint defined in (5) we have:

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

Now replace (13) into the equation above and obtain:

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

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

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

Using the definitions (24), (25), and (28):

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

Hence,

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

Now using definitions in (28) and (29) 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:

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

Observe $\dfrac{r_t(1+\tau^c_{t-1})}{1+\tau^c_t} = R_t - \dfrac{1+\tau^c_{t-1}}{1+\tau^c_t}$ and that $-\tau^c_t(\tilde{A}_t - \tilde{B}_t) + \tilde{B}_t = \hat{B}_t$ given by equation (30), which leads us:

$$G_t + R_t\tilde{B}_{t-1} + W_t \int_i (y^i_t l^i_t)^{1-\tau_t} \ell(di) + \left(R_t - \frac{1+\tau^c_{t-1}}{1+\tau^c_t}\right)(\tilde{A}_{t-1} - \tilde{B}_{t-1}) =$$
$$\frac{\tau^c_t(1+\tau^c_{t-1})}{1+\tau^c_t}(\tilde{A}_{t-1} - \tilde{B}_{t-1}) + F(K_{t-1}, L_t, s_t) + \hat{B}_t.$$

Hence, we have:

$$G_t + R_t\tilde{B}_{t-1} + W_t \int_i (y^i_t l^i_t)^{1-\tau_t} \ell(di) + (R_t - (1+\tau^c_{t-1}))(\tilde{A}_{t-1} - \tilde{B}_{t-1}) =$$
$$F(K_{t-1}, L_t, s_t) + \hat{B}_t.$$

Finally, using (30) in period $t-1$ (i.e., $\hat{B}_{t-1} = (1+\tau^c_{t-1})\tilde{B}_{t-1} - \tau^c_{t-1}\tilde{A}_{t-1}$) we get (32):

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

B First-order conditions of the individual Ramsey program

The Ramsey problem in (31) - (37) post-tax is given by:

$$\max_{\left(W_t, R_t, \tilde{w}_t, \tilde{r}_t, \tau^c_t, \tau^K_t, \tau_t, \kappa_t,, \hat{B}_t, G_t, K_t, L_t, (c^i_t, l^i_t, \tilde{a}^i_t, \nu^i_t)_i\right)_{t\geq 0}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_i \omega^i_t \left(u(c^i_t) - v(l^i_t) + u_G(G_t)\right) \ell(di)\right],$$

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

for all $i \in \mathcal{I}$: $c^i_t + \tilde{a}^i_t = W_t(y^i_t l^i_t)^{1-\tau_t} + R_t \tilde{a}^i_{t-1}$,

$$\tilde{a}^i_t \geq -\bar{\tilde{a}},\ \nu^i_t(\tilde{a}^i_t + \bar{\tilde{a}}) = 0,\ \nu^i_t \geq 0,$$

$$u'(c^i_t) = \beta \mathbb{E}_t \left[R_{t+1} u'(c^i_{t+1})\right] + \nu^i_t,$$

$$v'(l^i_t) = (1-\tau_t) W_t y^i_t (y^i_t l^i_t)^{-\tau_t} u'(c^i_t),$$

$$K_t + \hat{B}_t = \tilde{A}_t = \int_i \tilde{a}^i_t \ell(di),\ L_t = \int_i y^i_t l^i_t \ell\,(di)\,.$$

The instruments are: $\tilde{a}^i_t$, l^i_t, W_t, R_t, τ_t, $\hat{B}_t$, and G_t.

The Lagrangian can be written as:

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

where:

$$
c_t^i = -\tilde{a}_t^i + R_t \tilde{a}_{t-1}^i + W_t (y_t^i l_t^i)^{1-\tau_t}, \tag{B.2}
$$

$$
Y_t = \Big(\int_i \tilde{a}_{t-1}^i \ell(di) - \hat{B}_{t-1} \Big)^{\alpha} \Big(\int_i y_t^i l_t^i \ell\,(di) \Big)^{1-\alpha}. \tag{B.3}
$$

FOC with respect to savings choices. Deriving (B.1) with respect to $\tilde{a}_t^i$ yields:

$$
\begin{aligned}
0 = {} & \beta^t \int_j \omega_t^j u'(c_t^j) \frac{\partial c_t^j}{\partial \tilde{a}_t^i} \ell(dj) \\
& - \beta^t \int_j \left(\lambda_{c,t}^j - R_t \lambda_{c,t-1}^j \right) u''(c_t^j) \frac{\partial c_t^j}{\partial \tilde{a}_t^i} \ell(dj) \\
& + \beta^t (1-\tau_t) W_t \int_j \lambda_{l,t}^j (y_t^j)^{1-\tau_t} (l_t^j)^{-\tau_t} u''(c_t^j) \frac{\partial c_t^j}{\partial \tilde{a}_t^i} \ell(dj) \\
& + \beta^t \mu_t \left(\frac{\partial Y_t}{\partial \tilde{a}_t^i} - (R_t - 1 + \delta) \frac{\partial \tilde{A}_{t-1}}{\partial \tilde{a}_t^i} \right) \\
& + \beta^{t+1} \mathbb{E}_t \left[\int_j \omega_{t+1}^j u'(c_{t+1}^j) \frac{\partial c_{t+1}^j}{\partial \tilde{a}_t^i} \ell(dj) \right] \\
& - \beta^{t+1} \mathbb{E}_t \left[\int_j \left(\lambda_{c,t+1}^j - R_{t+1} \lambda_{c,t}^j \right) u''(c_{t+1}^j) \frac{\partial c_{t+1}^j}{\partial \tilde{a}_t^i} \ell(dj) \right] \\
& + \beta^{t+1} (1-\tau_{t+1}) W_{t+1} \mathbb{E}_t \left[\int_j \lambda_{l,t+1}^j (y_{t+1}^j)^{1-\tau_{t+1}} (l_{t+1}^j)^{-\tau_{t+1}} u''(c_{t+1}^j) \frac{\partial c_{t+1}^j}{\partial \tilde{a}_t^i} \ell(dj) \right] \\
& + \beta^{t+1} \mathbb{E}_t \left[\mu_{t+1} \left(\frac{\partial Y_{t+1}}{\partial \tilde{a}_t^i} - (R_{t+1} - 1 + \delta) \frac{\partial \tilde{A}_t}{\partial \tilde{a}_t^i} \right) \right].
\end{aligned}
$$

Denote:

$$
\psi_t^i := \omega_t^i u'(c_t^i) - \left(\lambda_{c,t}^i - R_t \lambda_{c,t-1}^i - \lambda_{l,t}^i (1-\tau_t) W_t (y_t^i)^{1-\tau_t} (l_t^i)^{-\tau_t} \right) u''(c_t^i). \tag{B.4}
$$

and get:

$$
\begin{aligned}
0 = {} & \int_j \psi_t^j \frac{\partial c_t^j}{\partial \tilde{a}_t^i} \ell(dj) + \beta \mathbb{E}_t \left[\int_j \psi_{t+1}^j \frac{\partial c_{t+1}^j}{\partial \tilde{a}_t^i} \ell(dj) \right] \\
& + \beta \mathbb{E}_t \left[\mu_{t+1} \left(\frac{\partial Y_{t+1}}{\partial \tilde{a}_t^i} - (R_{t+1} - 1 + \delta) \frac{\partial \tilde{A}_t}{\partial \tilde{a}_t^i} \right) \right].
\end{aligned}
$$

Using (B.2) and (B.3), we obtain:

$$\frac{\partial c_t^j}{\partial \tilde{a}_t^i} = -1_{i=j},$$
$$\frac{\partial \tilde{A}_t}{\partial \tilde{a}_t^i} = 1,$$
$$\frac{\partial c_{t+1}^j}{\partial \tilde{a}_t^i} = R_{t+1}1_{i=j},$$
$$\frac{\partial Y_{t+1}}{\partial \tilde{a}_t^i} = \alpha K_t^{\alpha-1}L_{t+1}^{1-\alpha}.$$

From which we deduce:

$$\begin{aligned}\psi_t^i &= \beta\mathbb{E}_t\left[R_{t+1}\psi_{t+1}^i\right] + \beta\mathbb{E}_t\left[\mu_{t+1}\left(\alpha K_t^{\alpha-1}L_{t+1}^{1-\alpha} - (R_{t+1}-1+\delta)\right)\right] \\ &= \beta\mathbb{E}_t\left[R_{t+1}\psi_{t+1}^i\right] + \beta\mathbb{E}_t\left[\mu_{t+1}\left(\tilde{r}_{t+1} - R_{t+1} + 1\right)\right].\end{aligned}$$

FOC with respect to labor supply. Deriving (B.1) with respect to l_t^i yields:

$$\begin{aligned}0 = &\int_j \psi_t^j \frac{\partial c_t^j}{\partial l_t^i}\ell(dj) - \omega_t^i v'(l_t^i) - \lambda_{l,t}^i v''(l_t^i) \\ &+ \lambda_{l,t}^i(1-\tau_t)(-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t-1}u'(c_t^i) - \mu_t(W_t(1-\tau_t)(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t} - \frac{\partial Y_t}{\partial l_t^i}).\end{aligned}$$

Using (B.2) and (B.3) we obtain respectively, $\frac{\partial c_t^j}{\partial l_t^i} = (1-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t}1_{i=j}$ and $\frac{\partial Y_t}{\partial l_t^i} = F_{L,t}y_t^i$. So:

$$\begin{aligned}\omega_t^i v'(l_t^i) + \lambda_{l,t}^i v''(l_t^i) = &(1-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t}\hat{\psi}_t^i \\ &+ \lambda_{l,t}^i(1-\tau_t)(-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t-1}u'(c_t^i) + \mu_t y_t^i F_{L,t},\end{aligned}$$

where $\hat{\psi}_t^i := \psi_t^i - \mu_t$. After some manipulation we get:

$$\frac{\omega_t^i v'(l_t^i) + \lambda_{l,t}^i v''(l_t^i)}{(1-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t}} = \hat{\psi}_t^i - \lambda_{l,t}^i\tau_t\frac{u'(c_t^i)}{l_t^i} + \mu_t\frac{F_{L,t}}{(1-\tau_t)W_t(y_t^i)^{-\tau_t}(l_t^i)^{-\tau_t}}.$$

Now define:

$$\psi_{l,t}^i := \omega_t^i v'(l_t^i) + \lambda_{l,t}^i v''(l_t^i).$$

Hence we have:

$$\psi_{l,t}^i = (1-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t}\hat{\psi}_t^i + \mu_t F_{L,t}y_t^i - (1-\tau_t)W_t(y_t^i)^{1-\tau_t}(l_t^i)^{-\tau_t}\lambda_{l,t}^i\tau_t\frac{u'(c_t^i)}{l_t^i}.$$

FOC with respect to the wage rate. Deriving (B.1) with respect to W_t yields:

$$\begin{aligned}0 = &\int_j\left(\psi_t^j\frac{\partial c_t^j}{\partial W_t} + \lambda_{l,t}^j(1-\tau_t)(y_t^j)^{1-\tau_t}(l_t^j)^{-\tau_t}u'(c_t^j)\right)\ell(dj) \\ &- \mu_t\int_j (y_t^j l_t^j)^{1-\tau_t}\ell(dj).\end{aligned}$$

Using (B.2) we have: $\frac{\partial c_t^j}{\partial W_t} = (y_t^j l_t^j)^{1-\tau_t}$ and so:

$$0 = \int_j (y_t^j l_t^j)^{1-\tau_t} \left(\hat{\psi}_t^j + \lambda_{l,t}^j (1-\tau_t) u'(c_t^j)/l_t^j \right) \ell(dj).$$

FOC with respect to the interest rate. Deriving (B.1) with respect to R_t yields:

$$\begin{aligned} 0 = & \int_j \left(\psi_t^j \frac{\partial c_t^j}{\partial R_t} + \lambda_{c,t-1}^j u'(c_t^j) \right) \ell(dj) \\ & - \mu_t \int_j \tilde{a}_{t-1}^j \ell(dj). \end{aligned}$$

Now using(B.2) we have $\frac{\partial c_t^j}{\partial R_t} = \tilde{a}_{t-1}^j$. Hence:

$$0 = \int_j \left(\hat{\psi}_t^j \tilde{a}_{t-1}^j + \lambda_{c,t-1}^j u'(c_t^j) \right) \ell(dj).$$

FOC with respect to progressivity. Deriving (B.1) with respect to τ_t yields:

$$\begin{aligned} 0 = & \int_j \psi_t^j \frac{\partial c_t^j}{\partial \tau_t} \ell(dj) \\ & + W_t \int_j \lambda_{l,t}^j \frac{\partial}{\partial \tau_t} \left((1-\tau_t)(y_t^j l_t^j)^{1-\tau_t} \right) (u'(c_t^j)/l_t^j) \ell(dj) \\ & - \mu_t W_t \int_j \frac{\partial}{\partial \tau_t} \left((y_t^j l_t^j)^{1-\tau_t} \right) \ell(dj). \end{aligned}$$

Using (B.2) we have $\frac{\partial c_t^j}{\partial \tau_t} = W_t \frac{\partial}{\partial \tau_t} \left((y_t^j l_t^j)^{1-\tau_t} \right)$, and so:

$$\begin{aligned} 0 = & \int_j \hat{\psi}_t^j \frac{\partial}{\partial \tau_t} \left((y_t^j l_t^j)^{1-\tau_t} \right) \ell(dj) - \int_j \lambda_{l,t}^j \left((y_t^j l_t^j)^{1-\tau_t} \right) (u'(c_t^j)/l_t^j) \ell(dj) + \\ & \int_j \lambda_{l,t}^j (1-\tau_t) \frac{\partial}{\partial \tau_t} \left((y_t^j l_t^j)^{1-\tau_t} \right) (u'(c_t^j)/l_t^j) \ell(dj). \end{aligned}$$

Now use $\frac{\partial}{\partial \tau_t} \left((y_t^j l_t^j)^{1-\tau_t} \right) = -\ln(y_t^j l_t^j)(y_t^j l_t^j)^{1-\tau_t}$ and obtain:

$$\begin{aligned} 0 = & \int_j (y_t^j l_t^j)^{1-\tau_t} (\hat{\psi}_t^j + \lambda_{l,t}^j (1-\tau_t)(u'(c_t^j)/l_t^j)) \ln(y_t^j l_t^j) \ell(dj) \\ & + \int_j \lambda_{l,t}^j \left((y_t^j l_t^j)^{1-\tau_t} \right) (u'(c_t^j)/l_t^j) \ell(dj). \end{aligned}$$

FOC with respect to public debt. Deriving (B.1) with respect to $\hat{B}_t$ yields:

$$0 = \mu_t - \beta \mathbb{E}_t \left[(1 - \delta - \frac{\partial Y_{t+1}}{\partial \hat{B}_t}) \mu_{t+1} \right],$$

where $\frac{\partial Y_{t+1}}{\partial \hat{B}_{t+1}} = -(\tilde{r}_{t+1} + \delta)$. Hence,:

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

FOC with respect to government spending. Deriving (B.1) with respect to G_t yields:

$$\mu_t = \int_j \omega_t^j u'_G(G_t) \ell(dj).$$

Summary of FOCs.

$$\begin{aligned}
\hat{\psi}_t^i &= \beta \mathbb{E}_t \left[R_{t+1} \hat{\psi}_{t+1}^i \right], \\
\psi_{l,t}^i &= (1-\tau_t) W_t (y_t^i)^{1-\tau_t} (l_t^i)^{-\tau_t} \hat{\psi}_t^i + \mu_t F_{L,t} y_t^i \\
&\quad - (1-\tau_t) W_t (y_t^i)^{1-\tau_t} (l_t^i)^{-\tau_t} \lambda_{l,t}^i \tau_t \frac{u'(c_t^i)}{l_t^i}, \\
0 &= \int_j (y_t^j l_t^j)^{1-\tau_t} \left(\hat{\psi}_t^j + \lambda_{l,t}^j (1-\tau_t) u'(c_t^j)/l_t^j \right) \ell(dj), \\
0 &= \int_j \left(\hat{\psi}_t^j \tilde{a}_{t-1}^j + \lambda_{c,t-1}^j u'(c_t^j) \right) \ell(dj), \\
0 &= \int_j (y_t^j l_t^j)^{1-\tau_t} (\hat{\psi}_t^j + \lambda_{l,t}^j (1-\tau_t)(u'(c_t^j)/l_t^j)) \ln(y_t^j l_t^j) \ell(dj) \\
&\quad + \int_j \lambda_{l,t}^j \left((y_t^j l_t^j)^{1-\tau_t} \right) (u'(c_t^j)/l_t^j) \ell(dj), \\
\mu_t &= \beta \mathbb{E}_t \left[(1+\tilde{r}_{t+1}) \mu_{t+1} \right], \\
\mu_t &= \int_j \omega_t^j u'_G(G_t) \ell(dj).
\end{aligned}$$

C Truncating the model

The Ramsey problem of Section 5.1 involves a joint distribution across wealth and Lagrange multipliers. This high-dimensional object raises a number of difficulties for the program resolution. For instance, this joint distribution affects the planner's instruments in a non-obvious way, which makes the methods based on perturbation of a well-identified steady-state not usable for solving such problems (as Reiter 2009, Boppart et al. 2018, Bayer et al. 2019 or Auclert et al. 2019).

The solution we provide here builds on LeGrand & Ragot (2022). It allows us to compute the steady-state allocation and to derive a finite number of equations that can simulate by perturbation the dynamics of the Ramsey program for small aggregate shocks.[21] The intuition of the method can be summarized as follows. We build an aggregation of the Bewley model (thus for a given fiscal policy and no aggregate shock) in which agents with the same history over last N periods (where N is a fixed horizon) are aggregated into a unique "agent". The method

[21] LeGrand & Ragot (2022) contain additional results regarding convergence properties and accuracy. To the best of our knowledge, Bhandari et al. (2020) is the only other method to compute Ramsey allocations with aggregate shocks in general cases – even though it can be implemented only when credit constraints are not binding in equilibrium.

implies that the "aggregate" agent is endowed with the average wealth of all individual agents with this N-period history. The wealth heterogeneity among these individual agents is captured in our aggregate model through additional parameters – that will be called "ξs".[22]

The aggregation method yields a so-called truncated model, which thanks to the "ξs" is an exact aggregation of the underlying Bewley model in the absence of aggregate shocks. In the presence of aggregate shocks, we can simulate the truncated model using standard perturbation techniques. The truncated model also allows us to solve for the Ramsey program.

C.1 The truncated model

Let $N \geq 0$ be a truncation length. The key step of the aggregation consists in assigning to all agents sharing the same idiosyncratic history over the last $N \geq 0$ periods the same wealth and the same allocation. Such a N-period history will be said to be a truncated history and for a history $y^t = \{y_0, \ldots, y_{t-N}, y_{t-N+1}, \ldots, y_{t-1}, y_t\}$, this corresponds to the N-length vector denoted $y^N = \{y^N_{-N+1}, \ldots, y^N_{-1}, y^N_0\}$. To sum up we can represent the truncated history of an agent i whose idiosyncratic history is y^t as:

$$y^t = \{\underbrace{y_0, \ldots, y_{t-N-2}, y_{t-N-1}, y_{t-N}}_{\xi_{y^N}}, \underbrace{y_{t-N+1}, \ldots, y_{t-1}, y_t}_{=y^N}\}$$
$$:= \{\underbrace{\ldots, y^t_{-N-2}, y^t_{-N-1}, y^t_{-N}}_{\xi_{y^N}}, \underbrace{y^t_{-N+1}, \ldots, y^t_{-1}, y^t_0}_{=y^N}\},$$

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

First of all 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 the 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 $\Pi_{t,\tilde{y}^N y^N}$ (with $\sum_{y^N \mathcal{Y}^N} \Pi_{t,\tilde{y}^N y^N} = 1$) and can be computed from the transition probabilities for the productivity process as:

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

With those elements now we can compute the share of agents with truncated history y^N as S_{t,y^N}. This element wil be $S_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} S_{t-1,\tilde{y}^N} \Pi_{t,\tilde{y}^N y^N}$, where the initial shares $(S_{-1,y^N})_{y^N \in \mathcal{Y}^N}$ is 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 (be it for consumption, savings, or labor supply) of the group of agents sharing the same truncated history. We consider a generic variable, denoted by $X_t(y^t, s^t)$ and we denote by X_{t,y^N} the average

[22] Werning (2015) also proposes an aggregation method, where the heterogeneity is captured by a change in the discount factor, while our method involve the introduction of explicit correction factors, the ξs. Using the ξs enables us to solve for the Ramsey program, as all agents are endowed with the same discount factor.

quantity of X assigned to truncated history y^N. Formally:

$$X_{t,y^N} = \frac{1}{S_{t,y^N}} \sum_{y^t \in \mathcal{Y}^{t+1} \big| (y^t_{-N+1},\dots,y^t_{-1},y^t_0)=y^N} X_t(y^t,s^t)\theta_t(y^t), \tag{C.1}$$

where we remind that $\theta_t(y^t)$ is the measure of agents with history y^t. Definition (C.1) 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 (C.1) to beginning-of-period wealth involves accounting that agents with truncated history y^N at t may have come from various truncated history at $t-1$. Indeed, this variable consists of the wealth of all agents having history y^N in period t and 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:

$$\tilde{\tilde{a}}_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \tilde{a}_{t-1,\tilde{y}^N}. \tag{C.2}$$

Now, in what follows we will define the various "ξs". Observe:

$$\sum_{y^t \in \mathcal{Y}^{t+1} \big| (y^t_{-N+1},\dots,y^t_{-1},y^t_0)=y^N} u(c_t\left(y^t\right)) = \xi^{u,0}_{y^N} u(\sum_{y^t \in \mathcal{Y}^{t+1} \big| (y^t_{-N+1},\dots,y^t_{-1},y^t_0)=y^N} c_t\left(y^t\right)),$$

which can be write compactly as:

$$\sum_{y^N \in \mathcal{Y}^N} u(c^i_t) = \xi^{u,0}_{y^N} u(c_{t,y^N}). \tag{C.3}$$

The same procedure applied to the other variables in the Ramsey problem in (31) - (37) generates:

$$\sum_{y^N \in \mathcal{Y}^N} v(l^i_t) := \xi^{v,0}_{y^N} v(l_{t,y^N}), \tag{C.4}$$

$$\sum_{y^N \in \mathcal{Y}^N} u'(c^i_t) := \xi^{u,1}_{y^N} u'(c_{t,y^N}), \tag{C.5}$$

$$\sum_{y^N \in \mathcal{Y}^N} (y^i_t l^i_t)^{1-\tau_t} := \xi^{y}_{y^N} (y^N_0 l_{t,y^N})^{1-\tau_t}. \tag{C.6}$$

We can now proceed with the aggregation of the full-fledged model. First, the aggregation of individual budget constraints (26) yields the following equation:

$$c_{t,y^N} + \tilde{a}_{t,y^N} = W_t \xi^{y}_{y^N} (l_{t,y^N} y^N_0)^{1-\tau_t} + R_t \tilde{\tilde{a}}_{t,y^N}, \text{ for } y^N \in \mathcal{Y}^N. \tag{C.7}$$

The aggregation of Euler equations for consumption (35) and labor (36) yields:

$$\xi^{u,E}_{y^N} u'(c_{t,y^N}) = \beta \mathbb{E}_t \left[R_{t+1} \sum_{\tilde{y}^N \in \mathcal{Y}^N} \Pi_{t+1,y^N \tilde{y}^N} \xi^{u,E}_{\tilde{y}^N} u'(c_{t+1,\tilde{y}^N}) \right] + \nu_{t,y^N}, \tag{C.8}$$

$$\xi^{v,1}_{y^N} v'(l_{t,y^N}) := (1-\tau_t) W_t \xi^{y}_{y^N} (l_{t,y^N} y^N_0)^{1-\tau_t} \xi^{u,1}_{y^N} (u'(c_{t,y^N})/l_{t,y^N}), \tag{C.9}$$

where the coefficients $(\xi^{u,E}_{y^N})_{y^N}$ for the consumption Euler equations guarantee that aggregate Euler equations yields 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 each history) verify the steady state consumption Euler equation (C.8). These coefficients are needed because Euler equations involve non-linear marginal utilities. The same idea for the coefficients $(\xi^{v,1}_{y^N})_{y^N}$ for the Euler equation for labor.

Finally, market clearing conditions can be expressed as:

$$K_t + \hat{B}_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \tilde{a}_{t,y^N}, \quad L_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} y_{y^N} l_{t,y^N}. \tag{C.10}$$

Equations (C.7)–(C.10) exactly characterizes 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. In the Appendix C.3 we have the derivation of the program formulation in the Projected model as well as the equations that characterize the social weights.

Steady-state and computation of the ξs. Steady-state allocations allow us to compute the parameters ξs as follows. Indeed, we can compute policy functions and wealth distribution of the Bewley model, as well as identify credit-constrained histories. Aggregation equations (C.1) and (C.2) 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 (C.8) to deduce the preference parameters $(\xi^{u,E}_{y^N})_{y^N}$. The other ξs are computed as explicit by equations (C.3), (C.4), (C.5), (C.6), and (C.9).

The truncated model in the presence of aggregate shocks. We state two further assumptions that will enables us to use our truncation method in the presence of aggregate shocks. This results in the so-called truncated model.

Assumption A. *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 $\mathcal{C} \subset \mathcal{Y}^N$, 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 of the truncated models are identical. Second, when the truncation length N becomes increasingly long, truncated allocations (in the presence of aggregate shocks) can be proved 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.

C.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.[23] Let $(\omega_y)_{y \in \mathcal{Y}}$ be the

[23]Our method involves deriving the first-order conditions of the truncated model, and not to truncate the first-order conditions of the full-fledged Ramsey model. This ensures numerical stability, as by construction the truncated model is "well-defined" for the fiscal policy under consideration.

set Pareto weights associated to each productivity level. The Ramsey program in the truncated economy can be written as follows:

$$\max_{\left(W_t,R_t,\tilde{w}_t,\tilde{r}_t,\tau_t^c,\tau_t^K,\tau_t,\kappa_t,\hat{B}_t,G_t,K_t,L_t,(c_{t,y^N},l_{t,y^N},\tilde{a}_{t,y^N},\nu_{t,y^N})_{y^N}\right)_{t\geq 0}} W_0, \tag{C.11}$$

where $W_0 := \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\sum_{y^N\in\mathcal{Y}^N}S_{t,y^N}\omega_{y^N}(\xi_{y^N}^{u,0}u(c_{t,y^N})-\xi_{y^N}^{v,0}v(l_{t,y^N})+u_G(G_t))\right]$ and subject to aggregate Euler equations (C.8) and (C.9), aggregate budget constraint (C.7), aggregate market clearing conditions (C.10), credit constraints $\tilde{a}_{t,y^N}\geq -\bar{\tilde{a}}$, as well as the governmental budget constraint (32), already present in the full-fledged Ramsey program.

Formally, the program can be written as:

$$\max_{\left(W_t,R_t,\tilde{w}_t,\tilde{r}_t,\tau_t^c,\tau_t^K,\tau_t,\kappa_t,\hat{B}_t,G_t,K_t,L_t,(c_{t,y^N},l_{t,y^N},\tilde{a}_{t,y^N},\nu_{t,y^N})_{y^N}\right)_{t\geq 0}} W_0, \tag{C.12}$$

$$G_t + W_t\sum_{y^N\in\mathcal{Y}^N}\xi_{y^N}^y S_{t,y^N}(l_{t,y^N}y_{y^N})^{1-\tau_t} + (R_t-1)\tilde{A}_{t-1} + \hat{B}_{t-1} = F(K_{t-1},L_t,s_t) + \hat{B}_t. \tag{C.13}$$

$$\text{for all } y^N\in\mathcal{Y}^N\text{: } c_{t,y^N} + \tilde{a}_{t,y^N} = W_t\xi_{y^N}^y(l_{t,y^N}y_0^N)^{1-\tau_t} + R_t\tilde{\tilde{a}}_{t,y^N}, \tag{C.14}$$

$$\tilde{a}_{t,y^N}\geq -\bar{\tilde{a}},\ \nu_{t,y^N}(\tilde{a}_{t,y^N}+\bar{\tilde{a}})=0,\ \nu_{t,y^N}\geq 0, \tag{C.15}$$

$$\xi_{y^N}^{u,E}u'(c_{t,y^N}) = \beta\mathbb{E}_t\left[R_{t+1}\sum_{\tilde{y}^N\in\mathcal{Y}^N}\Pi_{t+1,y^N\tilde{y}^N}\xi_{\tilde{y}^N}^{u,E}u'(c_{t+1,\tilde{y}^N})\right]+\nu_{t,y^N}, \tag{C.16}$$

$$\xi_{y^N}^{v,1}v'(l_{t,y^N}) := (1-\tau_t)W_t\xi_{y^N}^y(l_{t,y^N}y_0^N)^{1-\tau_t}\xi_{y^N}^{u,1}(u'(c_{t,y^N})/l_{t,y^N}), \tag{C.17}$$

$$K_t + \hat{B}_t = \sum_{y^N\in\mathcal{Y}^N}S_{t,y^N}\tilde{a}_{t,y^N},\quad L_t = \sum_{y^N\in\mathcal{Y}^N}S_{t,y^N}y_{y^N}l_{t,y^N}. \tag{C.18}$$

Computing the Pareto weights $(\omega_y)_{y\in\mathcal{Y}}$. The key contribution of our paper involves estimating the Pareto weights that corresponds to different fiscal systems. More precisely, we follow the methodology of the inverse optimal taxation literature (see Bargain & Keane 2010, Bourguignon & Amadeo 2015, Heathcote & Tsujiyama 2021, Chang et al. 2018, among others) and estimate Pareto weights for the first-order conditions of the Ramsey program in (31) - (37) to hold at the steady-state for the fiscal system under consideration (France or US in our quantitative exercise of Section 6). However, as we explain in Section 4, the estimation problem is under identified. We use the method of equation (22), which is however simplified in the truncated model, which feature limited heterogeneity. Formally, equation equation (22) becomes: $\omega_y = \arg\min_{(\tilde{\omega}_y)}\theta(y)\left\|(\tilde{\omega}_y)_y-\mathbf{1}\right\|_2$ subject to $\sum_y\theta(y)\tilde{\omega}_y=1$ and such that planner's first-order conditions hold. Our method also allows to estimate the Pareto weights for each history. In this case notice the same problem as argued in Section 4 applies, since we have Y^N Pareto weights (where $Y = Card\,\mathcal{Y}$) and few constraints. To estimate the pareto weights for each history we solve the following problem $\omega_{y^N} = \arg\min_{(\tilde{\omega}_{y^N})}S_{y^N}\left\|(\tilde{\omega}_{y^N})_{y^N}-\mathbf{1}\right\|_2$ subject to $\sum_y S_{y^N}\tilde{\omega}_{y^N}=1$.[24] Notice that the only difference between the two approaches is that in the second one we will have $(\omega_{y^N})_{y^N}$ different for each $y^N\in\mathcal{Y}^N$, whereas in the first approach $(\omega_{y^N})_{y^N} = (\omega_{\tilde{y}^N})_{\tilde{y}^N}$ whenever $y_0^N = \tilde{y}_0^N$, i.e., everytime the productivity level in the first period of the truncation associated with the history y^N is the same as the productivity level associated with the history $\tilde{y}^N$ with

[24] In the previous expression, $\|\cdot\|_2$ denotes the Euclidean norm, $(\omega_{y^N})_{y^N}$ is the vector of Pareto weights, and $\mathbf{1}$ is the vector of ones.

$y^N \neq \tilde{y}^N$.

Thanks to the limited heterogeneity both computations above pin down to simple matrix algebra. In what follows we explain how to obtain those estimates.

Factorization. Using the Lagrangian approach developed by Marcet & Marimon (2019) we obtain:

$$\begin{aligned} J =& \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{y^N \in \mathcal{Y}^N} \Bigg[S_{t,y^N} \Bigg(\omega_{y^N} (\xi_{y^N}^{u,0} u(c_{t,y^N}) - \xi_{y^N}^{v,0} v(l_{t,y^N}) + u_G(G_t)) \\ & - (\lambda_{c,t,y^N} - \tilde{\lambda}_{c,t,y^N} R_t) \xi_{y^N}^{u,E} u'(c_{t,y^N}) \\ & - \lambda_{l,t,y^N} (\xi_{y^N}^{v,1} v'(l_{t,y^N}) - (1-\tau_t) W_t (y_{y^N})^{1-\tau_t} \xi_{y^N}^{y} (l_{t,y^N})^{-\tau_t} \xi_{y^N}^{u,1} u'(c_{t,y^N})) \Bigg) \Bigg]. \end{aligned} \tag{C.19}$$

subject to

$$G_t + W_t \sum_{y^N \in \mathcal{Y}^N} \xi_{y^N}^{y} S_{t,y^N} (l_{t,y^N} y_{y^N})^{1-\tau_t} + (R_t - 1)\tilde{A}_{t-1} + \hat{B}_{t-1} = F(K_{t-1}, L_t, s_t) + \hat{B}_t.$$

$$\begin{aligned} &\text{for all } y^N \in \mathcal{Y}^N\text{: } c_{t,y^N} + \tilde{a}_{t,y^N} = W_t \xi_{y^N}^{y} (l_{t,y^N} y_0^N)^{1-\tau_t} + R_t \tilde{\tilde{a}}_{t,y^N}, \\ &\qquad \tilde{a}_{t,y^N} \geq -\tilde{\bar{a}},\ \nu_{t,y^N} (\tilde{a}_{t,y^N} + \tilde{\bar{a}}) = 0,\ \nu_{t,y^N} \geq 0 \\ &\qquad K_t + \hat{B}_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \tilde{a}_{t,y^N}, \quad L_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} y_{y^N} l_{t,y^N}. \end{aligned}$$

First order conditions. By defining the social value of liquidity of history y^N similarly as (38) we obtain:

$$\begin{aligned} \hat{\psi}_{t,y^N} =& \omega_{y^N} \xi_{y^N}^{u,0} u'(c_{t,y^N}) - \left(\lambda_{c,t,y^N} \xi_{y^N}^{u,E} - R_t \tilde{\lambda}_{c,t,y^N} \xi_{y^N}^{u,E} - \lambda_{l,t,y^N} \xi_{y^N}^{y} (1-\tau_t) W_t (y_0^N)^{1-\tau_t} l_{t,y^N}^{-\tau_t} \xi_{y^N}^{u,1} \right) u''(c_{t,y^N}) \\ & - \mu_t. \end{aligned} \tag{C.20}$$

In what follows we show the FOC of the the problem in (C.19).

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

$$\hat{\psi}_{t,y^N} = \beta \mathbb{E}_t \left[R_{t+1} \sum_{\tilde{y}^N \in \mathcal{Y}^N} \Pi_{t,y^N \tilde{y}^N} \hat{\psi}_{t+1,\tilde{y}^N} \right] \text{ if } \nu_{y^N} = 0 \text{ and } \lambda_{c,t,y^N} = 0 \text{ otherwise.} \tag{C.21}$$

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

$$\begin{aligned} \frac{\omega_{y^N} \xi_{y^N}^{v,0} v'(l_{t,y^N}) + \lambda_{l,t,y^N} \xi_{y^N}^{v,1} v''(l_{t,y^N})}{(1-\tau_t) W_t \xi_{y^N}^{y} (y_0^N)^{1-\tau_t} l_{t,y^N}^{-\tau_t}} =& \hat{\psi}_{t,y^N} - \lambda_{l,t,y^N} \tau_t \xi_{y^N}^{u,1} (u'(c_{t,y^N})/l_{t,y^N}) \\ & + \mu_t (1-\alpha) \frac{Y_t}{(1-\tau_t) W_t \xi_{y^N}^{y} (y_0^N)^{-\tau_t} l_{t,y^N}^{-\tau_t} L_t}. \end{aligned} \tag{C.22}$$

FOC with respect to W_t:

$$\sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \xi^y_{y^N} (l_{t,y^N} y_{y^N})^{1-\tau_t} \left(\hat{\psi}_{t,y^N} + \lambda_{l,t,y^N} (1-\tau_t) \xi^{u,1}_{y^N} (u'(c_{t,y^N})/l_{t,y^N}) \right) = 0. \tag{C.23}$$

FOC with respect to R_t:

$$\sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \left(\hat{\psi}_{t,y^N} \tilde{\tilde{a}}_{t,y^N} + \tilde{\lambda}_{c,t,y^N} \xi^{u,E}_{y^N} u'(c_{t,y^N}) \right) = 0. \tag{C.24}$$

FOC with respect to τ_t:

$$\sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \left(\hat{\psi}_{t,y^N} + \lambda_{l,t,y^N} (1-\tau_t) \xi^{u,1}_{y^N} (u'(c_{t,y^N})/l_{t,y^N}) \right) \ln \left(l_{t,y^N} y_{y^N} \right) \xi^y_{y^N} (l_{t,y^N} y_{y^N})^{1-\tau_t}$$
$$= - \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \lambda_{l,t,y^N} \xi^y_{y^N} (l_{t,y^N} y_{y^N})^{1-\tau_t} \xi^{u,1}_{y^N} (u'(c_{t,y^N})/l_{t,y^N}). \tag{C.25}$$

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

$$\mu_t = \beta \mathbb{E} \left[\mu_{t+1} \left(1 + \alpha \frac{Y_{t+1}}{K_t} - \delta \right) \right]. \tag{C.26}$$

FOC with respect to G_t:

$$\mu_t = \sum_{y^N \in \mathcal{Y}^N} S_{t,y^N} \omega_{y^N} u'_G(G_t). \tag{C.27}$$

The equations in (C.13) - (C.18) must be satisfied, as well as the following individual equations:

$$\tilde{\lambda}_{c,t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \lambda_{c,t-1,\tilde{y}^N}, \tag{C.28}$$

$$\tilde{a}_{t,y^N} \geq 0 \text{ and } \tilde{\tilde{a}}_{t,y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} \frac{S_{t-1,\tilde{y}^N}}{S_{t,y^N}} \Pi_{t,\tilde{y}^N y^N} \tilde{a}_{t-1,\tilde{y}^N}. \tag{C.29}$$

C.3 Matrix expression

In this section, we provide closed-form formulas for preference multipliers ξs (Section C.1) and the Pareto 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. It also provides an efficient solution to compute the values for the coefficients (ξ_{y^N}) and (ω_{y^N}). The starting point is to observe that a history y^N can be seen as a N-length vector $\{y^N_{-N+1}, \ldots, y^N_0\}$ of elements of $\mathcal{Y}$. The number of histories is $N_{tot} = Y^N$. We can identify each history by an integer $k_{y^N} = 1, \dots, N_{tot}$:

$$k_{y^N} = \sum_{k=0}^{N-1} N_{tot}^{-N+1-k}(y_k - 1) + 1, \tag{C.30}$$

which corresponds to an enumeration in base Y.

C.3.1 A closed-form formula for the ξs

Let $\mathbf{S}$ be the N_{tot}-vector of steady-state history sizes that is defined as $\mathbf{S} = (S_{k_{y^N}})_{k_{y^N}=1,\dots,N_{tot}}$, by stacking history sizes for all histories using the enumeration given by (C.30). Similarly, let $\tilde{\mathbf{a}}$, $\mathbf{c}$, $\boldsymbol{\ell}$, $\boldsymbol{\nu}$, $\mathbf{u}'(\mathbf{c})$, $\mathbf{v}'(\mathbf{l})$ $\mathbf{u}''(\mathbf{c})$, $\mathbf{v}''(\mathbf{l})$ 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 (C.1). We also define:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_Y \end{bmatrix} \otimes 1_B,$$

where 1_B is a vector of 1 of length Y^{N-1}. We define $\mathbb{P}$ as 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. We similarly define $\mathbb{P}^c = \mathbf{I} - \mathbb{P}$, where $\mathbf{I}$ is the $(N_{tot} \times N_{tot})$-identity matrix. Noting Π as the transition matrix across histories, we obtain the following steady-state relationships.

C.3.2 Matrix expressions for the definition of ξs

Writing $S_{y^N} = \sum_{\tilde{y}^N \in \mathcal{Y}^N} S_{\tilde{y}^N} \Pi_{\tilde{y}^N y^N}$, (C.14) and (C.15) in matrix form yields, respectively:

$$\boldsymbol{S} = \Pi \boldsymbol{S}, \tag{C.31}$$

$$\boldsymbol{S} \circ \boldsymbol{c} + \boldsymbol{S} \circ \tilde{\boldsymbol{a}} = R\Pi\left(\boldsymbol{S} \circ \tilde{\boldsymbol{a}}\right) + W\boldsymbol{S} \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}, \tag{C.32}$$

$$(\boldsymbol{I} - \boldsymbol{P})\tilde{\boldsymbol{a}} = \boldsymbol{0}_{N_{tot}\times 1}. \tag{C.33}$$

The Euler equation for consumption in (C.16) 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 matrix $\Pi^\top$ (the transpose of Π) is used to make expectations about next period histories. Or 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 $\boldsymbol{D}$ stands for the diagonal matrix. Finally:

$$\boldsymbol{\xi}^{u,E} = \left[\left(\boldsymbol{I} - \beta R \boldsymbol{\Pi}^\top\right)\boldsymbol{D}_{u'(\boldsymbol{c})}\right]^{-1} \boldsymbol{\nu}. \tag{C.34}$$

From the Euler equation for labor in (C.17) we obtain:

$$\boldsymbol{\xi}^{v,1} = (1-\tau)W(\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \boldsymbol{\xi}^y \circ \boldsymbol{\xi}^{u,1} \circ u'(\boldsymbol{c})./(\boldsymbol{l} \circ v'(\boldsymbol{l})). \tag{C.35}$$

Now from the aggregation equations in (C.3) -(C.6) we get:

$$\boldsymbol{\xi}^{u,0} = \frac{\sum_{y^N \in \mathcal{Y}^N} u(c_t^i)}{u(c_{t,y^N})}, \tag{C.36}$$

$$\boldsymbol{\xi}^{v,0} = \frac{\sum_{y^N \in \mathcal{Y}^N} v(l_t^i)}{v(l_{t,y^N})}, \tag{C.37}$$

$$\boldsymbol{\xi}^{u,1} = \frac{\sum_{y^N \in \mathcal{Y}^N} u'(c_t^i)}{u'(c_{t,y^N})}, \tag{C.38}$$

$$\boldsymbol{\xi}^{y} = \frac{\sum_{y^N \in \mathcal{Y}^N} (y_t^i l_t^i)^{1-\tau}}{\left(y_0^N l_{t,y^N}\right)^{1-\tau}}. \tag{C.39}$$

Finally define the following variables:

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

C.3.3 Matrix expressions for the FOCs

The expressions for the FOC from (C.20) to (C.26) becomes, respectively:

$$\hat{\boldsymbol{\psi}} = \boldsymbol{\omega} \circ \boldsymbol{\xi}^{u,0} \circ u'(\boldsymbol{c}) - \left(\boldsymbol{\lambda}_c \circ \boldsymbol{\xi}^{u,E} - R\tilde{\boldsymbol{\lambda}}_c \circ \boldsymbol{\xi}^{u,E} - (1-\tau)W\boldsymbol{\lambda}_l \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ (\boldsymbol{\xi}^{u,1}./\boldsymbol{l})\right) \circ u''(\boldsymbol{c}) - \mu\mathbf{1}.$$

$$\boldsymbol{P}\hat{\boldsymbol{\psi}} = \beta R \boldsymbol{P} \Pi^\top \hat{\boldsymbol{\psi}},$$
$$(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\lambda}_c = 0.$$

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

$$\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \times (\boldsymbol{S} \circ \hat{\boldsymbol{\psi}}) = -(1-\tau)\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \boldsymbol{\xi}^{u,1} \circ (u'(\boldsymbol{c})./\boldsymbol{l})\right)^\top \times (\boldsymbol{S} \circ \boldsymbol{\lambda}_l).$$

$$\tilde{\tilde{\boldsymbol{a}}}^\top \times (\boldsymbol{S} \circ \hat{\boldsymbol{\psi}}) = -\left(\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c})\right)^\top \times \left(\boldsymbol{S} \circ \tilde{\boldsymbol{\lambda}}_c\right).$$

$$\left(\ln(\boldsymbol{y} \circ \boldsymbol{l}) \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \times (\boldsymbol{S} \circ \hat{\boldsymbol{\psi}}) =$$
$$-\left((\mathbf{1} + (1-\tau)\ln(\boldsymbol{y} \circ \boldsymbol{l})) \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \boldsymbol{\xi}^{u,1} \circ u'(\boldsymbol{c})./\boldsymbol{l}\right)^\top \times (\boldsymbol{S} \circ \boldsymbol{\lambda}_l).$$

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

Define the following variables:

$$
\begin{aligned}
\bar{\lambda}^l &:= \mathbf{S} \circ \lambda^l,\\
\bar{\psi} &:= \mathbf{S} \circ \hat{\psi},\\
\bar{\Pi} &:= \mathbf{S} \circ \Pi^\top \circ (1/\mathbf{S}),\\
\bar{\omega} &:= \mathbf{S} \circ \omega,\\
\bar{\lambda}_c &:= \mathbf{S} \circ \lambda_c,
\end{aligned}
$$

and notice $\mathbf{S} \circ \tilde{\lambda}_c = \Pi \bar{\boldsymbol{\lambda}}_c$. This definition makes the FOC from (C.20) to (C.26) be:

$$
\bar{\boldsymbol{\psi}} = \bar{\boldsymbol{\omega}} \circ \boldsymbol{\xi}^{u,0} \circ u'(\boldsymbol{c}) - \left(\bar{\boldsymbol{\lambda}}_c \circ \boldsymbol{\xi}^{u,E} - R\boldsymbol{\Pi}\bar{\boldsymbol{\lambda}}_c \circ \boldsymbol{\xi}^{u,E} - (1-\tau)W\bar{\boldsymbol{\lambda}}_l \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1}\right) \circ u''(\boldsymbol{c}) - \mu \mathbf{S}. \tag{C.40}
$$

$$
\boldsymbol{P}\bar{\boldsymbol{\psi}} = \beta R \boldsymbol{P}\bar{\boldsymbol{\Pi}}\bar{\boldsymbol{\psi}}, \tag{C.41}
$$

$$
(\boldsymbol{I} - \boldsymbol{P})\bar{\boldsymbol{\lambda}}_c = 0. \tag{C.42}
$$

$$
\bar{\boldsymbol{\omega}} \circ \tilde{\boldsymbol{\xi}}^{v,0} \circ v'(\boldsymbol{l}) + \bar{\boldsymbol{\lambda}}_l \circ \tilde{\boldsymbol{\xi}}^{v,1} \circ v''(\boldsymbol{l}) = \bar{\boldsymbol{\psi}} - \tau \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c}) \circ \bar{\boldsymbol{\lambda}}_l + \mu F_L \boldsymbol{S}./((1-\tau)W\boldsymbol{\xi}^y \circ \boldsymbol{y}^{-\tau} \circ \boldsymbol{l}^{-\tau}). \tag{C.43}
$$

$$
\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \bar{\boldsymbol{\psi}} = -(1-\tau)\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^\top \bar{\boldsymbol{\lambda}}_l. \tag{C.44}
$$

$$
\tilde{\boldsymbol{a}}^\top \bar{\boldsymbol{\psi}} = -\left(\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c})\right)^\top \Pi \bar{\boldsymbol{\lambda}}_c. \tag{C.45}
$$

$$
\left(\ln(\boldsymbol{y} \circ \boldsymbol{l}) \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \bar{\boldsymbol{\psi}} = -\left((\mathbf{1} + (1-\tau)\ln(\boldsymbol{y} \circ \boldsymbol{l})) \circ \boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^\top \bar{\boldsymbol{\lambda}}_l. \tag{C.46}
$$

This is system in $\bar{\psi}, \bar{\omega}, \bar{\lambda}_c, \bar{\lambda}_l, \mu$, with the equivalent of 5 equations .

C.3.4 Solving the system

Equation (C.43) yields:

$$
\begin{aligned}
\boldsymbol{D}_{\tilde{\xi}^{v,1} \circ v''(\boldsymbol{l}) + \tau \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})} \bar{\boldsymbol{\lambda}}_l &= \mu F_L \boldsymbol{S}./((1-\tau)W\boldsymbol{\xi}^y \circ \boldsymbol{y}^{-\tau} \circ \boldsymbol{l}^{-\tau}) + \bar{\boldsymbol{\psi}} - \boldsymbol{D}_{\tilde{\boldsymbol{\xi}}^{v,0} \circ v'(\boldsymbol{l})} \bar{\boldsymbol{\omega}},\\
\bar{\boldsymbol{\lambda}}_l &= \boldsymbol{M}_0 \bar{\boldsymbol{\omega}} + \boldsymbol{M}_1 \bar{\boldsymbol{\psi}} + \mu \boldsymbol{V}_0.
\end{aligned} \tag{C.47}
$$

with:

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

Then equation (C.40) implies:

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

$$\begin{aligned}\bar{\boldsymbol{\psi}} &= \boldsymbol{D}_{\boldsymbol{\xi}^{u,0}\circ u'(\boldsymbol{c})}\bar{\boldsymbol{\omega}} - \boldsymbol{D}_{\boldsymbol{\xi}^{u,E}\circ u''(\boldsymbol{c})}(\boldsymbol{I} - R\boldsymbol{\Pi})\bar{\boldsymbol{\lambda}}_c + (1-\tau)W\boldsymbol{D}_{\boldsymbol{\xi}^{y}\circ(\boldsymbol{y}\circ\boldsymbol{l})^{1-\tau}\circ\tilde{\boldsymbol{\xi}}^{u,1}\circ u''(\boldsymbol{c})}\bar{\boldsymbol{\lambda}}_l - \mu\mathbf{S} \\ &= \hat{\boldsymbol{M}}_0\bar{\boldsymbol{\omega}} + \hat{\boldsymbol{M}}_1\bar{\boldsymbol{\lambda}}_c + \hat{\boldsymbol{M}}_2\bar{\boldsymbol{\lambda}}_l - \mu\mathbf{S}. \end{aligned} \tag{C.48}$$

with:

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

So using (C.48) and (C.47):

$$\begin{aligned}\bar{\boldsymbol{\psi}} &= \hat{\boldsymbol{M}}_0\bar{\boldsymbol{\omega}} + \hat{\boldsymbol{M}}_1\bar{\boldsymbol{\lambda}}_c + \hat{\boldsymbol{M}}_2(\boldsymbol{M}_0\bar{\boldsymbol{\omega}} + \boldsymbol{M}_1\bar{\boldsymbol{\psi}} + \mu\boldsymbol{V}_0) - \mu\mathbf{S}, \\ \left(\boldsymbol{I} - \hat{\boldsymbol{M}}_2\boldsymbol{M}_1\right)\bar{\boldsymbol{\psi}} &= (\hat{\boldsymbol{M}}_0 + \hat{\boldsymbol{M}}_2\boldsymbol{M}_0)\bar{\boldsymbol{\omega}} + \hat{\boldsymbol{M}}_1\bar{\boldsymbol{\lambda}}_c + \mu(\hat{\boldsymbol{M}}_2\boldsymbol{V}_0 - \mathbf{S}).\end{aligned}$$

Define:

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

Then:

$$\bar{\boldsymbol{\psi}} = \boldsymbol{M}_3\bar{\boldsymbol{\omega}} + \boldsymbol{M}_4\bar{\boldsymbol{\lambda}}_c + \mu\boldsymbol{V}_1. \tag{C.49}$$

Use (C.41), (C.42), and (C.49):

$$\begin{aligned}(\boldsymbol{I} - \boldsymbol{P})\bar{\boldsymbol{\lambda}}_c + \boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})\bar{\boldsymbol{\psi}} &= 0, \\ \left((\boldsymbol{I} - \boldsymbol{P}) + \boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})\boldsymbol{M}_4\right)\bar{\boldsymbol{\lambda}}_c &= -\boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})\boldsymbol{M}_3\bar{\boldsymbol{\omega}} - \mu\boldsymbol{P}(I - \beta R\bar{\boldsymbol{\Pi}})\boldsymbol{V}_1.\end{aligned}$$

Define:

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

and get:

$$\bar{\boldsymbol{\lambda}}_c = \boldsymbol{M}_5\bar{\boldsymbol{\omega}} + \mu\boldsymbol{V}_2. \tag{C.50}$$

Then we use (C.49) and (C.50) into (C.45) and notice:

$$\begin{aligned}\tilde{\boldsymbol{a}}^{\top}(\boldsymbol{M}_3\bar{\boldsymbol{\omega}} + \boldsymbol{M}_4\bar{\boldsymbol{\lambda}}_c + \mu\boldsymbol{V}_1) + \left(\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c})\right)^{\top}\Pi\bar{\boldsymbol{\lambda}}_c &= 0, \\ \tilde{\boldsymbol{a}}^{\top}((\boldsymbol{M}_3 + \boldsymbol{M}_4\boldsymbol{M}_5)\bar{\boldsymbol{\omega}} + \mu(\boldsymbol{V}_1 + \boldsymbol{M}_4\boldsymbol{V}_2)) + (\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}))^{\top}\boldsymbol{\Pi}(\boldsymbol{M}_5\bar{\boldsymbol{\omega}} + \mu\boldsymbol{V}_2) &= 0, \\ \left(\tilde{\boldsymbol{a}}^{\top}(\boldsymbol{M}_3 + \boldsymbol{M}_4\boldsymbol{M}_5) + (\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}))^{\top}\boldsymbol{\Pi}\boldsymbol{M}_5\right)\bar{\boldsymbol{\omega}} + \mu(\tilde{\boldsymbol{a}}^{\top}\boldsymbol{V}_1 + \tilde{\boldsymbol{a}}^{\top}\boldsymbol{M}_4\boldsymbol{V}_2 + (\boldsymbol{\xi}^{u,E} \circ u'(\boldsymbol{c}))^{\top}\boldsymbol{\Pi}\boldsymbol{V}_2) &= 0.\end{aligned}$$

Define:

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

Then,

$$\mu = -\boldsymbol{L}_1\bar{\omega}. \tag{C.51}$$

We deduce that from (C.49) and (C.50):

$$\begin{aligned}
\bar{\boldsymbol{\lambda}}_c &= (\boldsymbol{M}_5 - \boldsymbol{V}_2\boldsymbol{L}_1)\bar{\boldsymbol{\omega}}, \\
\bar{\boldsymbol{\psi}} &= \boldsymbol{M}_3\bar{\boldsymbol{\omega}} + \boldsymbol{M}_4\bar{\boldsymbol{\lambda}}_c + \mu\boldsymbol{V}_1 \\
&= (\boldsymbol{M}_3 + \boldsymbol{M}_4(\boldsymbol{M}_5 - \boldsymbol{V}_2\boldsymbol{L}_1) - \boldsymbol{V}_1\boldsymbol{L}_1)\bar{\boldsymbol{\omega}} \\
&= \boldsymbol{M}_6\bar{\boldsymbol{\omega}},
\end{aligned}$$

with:

$$\boldsymbol{M}_6 := \boldsymbol{M}_3 + \boldsymbol{M}_4(\boldsymbol{M}_5 - \boldsymbol{V}_2\boldsymbol{L}_1) - \boldsymbol{V}_1\boldsymbol{L}_1.$$

and from (C.47):

$$\begin{aligned}
\bar{\boldsymbol{\lambda}}_l &= \hat{\boldsymbol{M}}_6\bar{\boldsymbol{\omega}}, \\
\text{where: } \hat{\boldsymbol{M}}_6 &:= \boldsymbol{M}_0 + \boldsymbol{M}_1\boldsymbol{M}_6 - \boldsymbol{V}_0\boldsymbol{L}_1,
\end{aligned}$$

Constructing the constraints. The constraint of equation (C.46) is:

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

or

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

or equivalently:

$$\boldsymbol{L}_2\bar{\boldsymbol{\omega}} = 0, \tag{C.52}$$

where:

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

The constraint (C.44) becomes:

$$\boldsymbol{L}_3\bar{\boldsymbol{\omega}} = 0, \tag{C.53}$$

where:

$$\boldsymbol{L}_3 := \left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau}\right)^\top \boldsymbol{M}_6 + (1-\tau)\left(\boldsymbol{\xi}^y \circ (\boldsymbol{y} \circ \boldsymbol{l})^{1-\tau} \circ \tilde{\boldsymbol{\xi}}^{u,1} \circ u'(\boldsymbol{c})\right)^\top \hat{\boldsymbol{M}}_6.$$

When agents derive utility from government spending equation (C.51) becomes:

$$\tilde{\boldsymbol{L}}_1 \bar{\boldsymbol{\omega}} = 0, \tag{C.54}$$

where $\tilde{\boldsymbol{L}}_1 := \boldsymbol{u}'_G(\boldsymbol{G_t})^\top + \boldsymbol{L}_1$.

C.3.5 Constrained weight

Now assume that there is a limited number of Pareto K weight, ω^s. Let $\boldsymbol{M}_7$ be a $N_{tot} \times K$ and $\boldsymbol{I}_{K\times K}$ be a $K \times K$ identity matrix. Then we have

$$\bar{\boldsymbol{\omega}} = \boldsymbol{D_S M_7 \omega^s},$$

where $\boldsymbol{D}_S$ is the diagonal of the the variable S_{y^N} and $\boldsymbol{M}_7$ will be a matrix identifying the productivity level of each bin.

Finally, the pareto weights for each productivity level are given as a solution of the following minimization problem:

$$\begin{aligned} \min_{\omega} \boldsymbol{\theta} \left\|\boldsymbol{\omega}^s - \boldsymbol{1}_K\right\|^2, \\ \text{s.t. } \boldsymbol{L}_2\boldsymbol{D_S M_7 \omega^s} = 0, \\ \boldsymbol{L}_3\boldsymbol{D_S M_7 \omega^s} = 0, \\ \boldsymbol{1}^\top\boldsymbol{D_S M_7 \omega^s} = 1. \end{aligned} \tag{C.55}$$

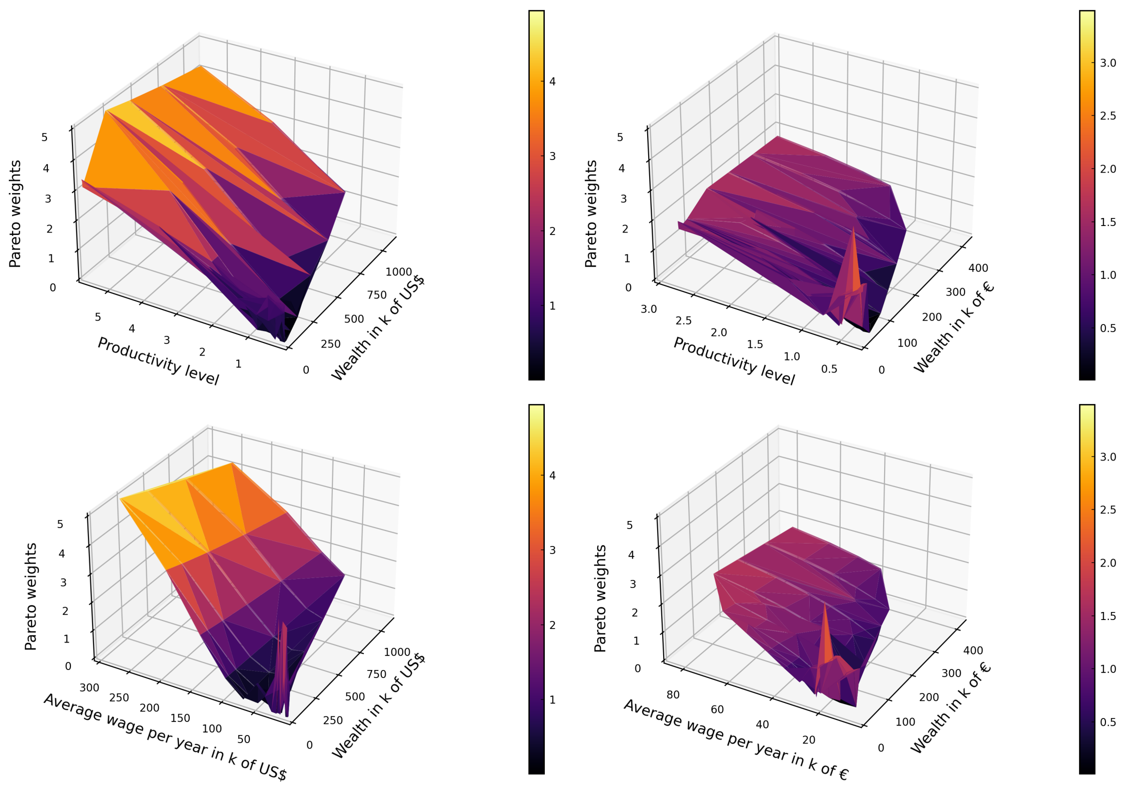
In case we want to estimate the pareto weights for each history we just need to set $\boldsymbol{M}_7 = 1$ and solve the problem below:

$$\begin{aligned} \min_{\omega} \boldsymbol{S} \left\|\boldsymbol{\omega}^s - \boldsymbol{1}_K\right\|^2, \\ \text{s.t. } \boldsymbol{L}_2\boldsymbol{D_S M_7 \omega^s} = 0, \\ \boldsymbol{L}_3\boldsymbol{D_S M_7 \omega^s} = 0, \\ \boldsymbol{1}^\top\boldsymbol{D_S M_7 \omega^s} = 1. \end{aligned} \tag{C.56}$$

D Pareto weights by history

We calculate the pareto weights by history in Figure 9, where we represent the weights for the US in panel (a) and for France in panel (b) to ease comparison. We do the same exercise by considering the average wage per year of each agent and wealth associated to each history in the national currency of the respective country. This is also represented in Figure 9.[25]

[25] For the US we obtain the average hourly earnings of all employees from the *FRED* for the year 2007. This value is US\$ 20.17 per hour. For France we use the value for 2006 obtained in *Eurostat*, which is also the average hourly earnings of all employees. The value for France for the year 2006 is € 13.23 per hour. We then obtain the normalized wage using $\tilde{w}_t = \dfrac{\overline{w}\sum_{y^N\in\mathcal{Y}^N} S_{t,y^N}\tilde{l}_{t,y^N}}{\sum_{y^N\in\mathcal{Y}^N}\xi^y_{y^N} S_{t,y^N}(\tilde{l}_{t,y^N}y_{y^N})^{1-\tau_t}}$, where $\overline{w}$ represents the average wage. By doing this procedure we can, hence, calculate the total wage per year as being $\tilde{w}_t(\tilde{l}_{t,y^N}y_{y^N})^{1-\tau}$, where $\tilde{l}_{t,y^N} := 8760 l_{t,y^N}$ with 8760 being the total hours in a year and l_{t,y^N} the fraction of time the agent in this economy spends working. To normalize the wealth we use the fact that $\dfrac{A_t}{Y_t} = \dfrac{K_t}{Y_t} + \dfrac{B_t}{Y_t}$. Notice, we can define $A_t = GDP^i_t(\dfrac{K_t}{Y_t} + \dfrac{B_t}{Y_t})$,



(a) United States (b) France

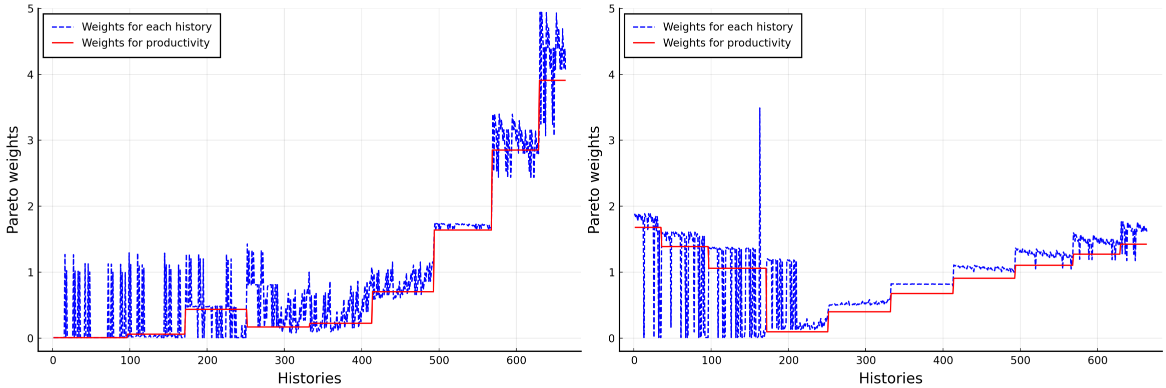
Figure 9: Pareto weights as a function of productivity (income per capita) and wealth for the US and France.

For the US, one can observe that the Pareto weights have a U-shape in the dimension of productivity/income (for low wealth levels) becoming more increasing as the wealth grows, and is slightly increasing in the dimension of wealth (for low productivity/income level) becoming more decreasing as the productivity/income level gets higher (in Appendix E we see the same result for a longer truncation), which means the planner puts more weight to low productivity agents as long as they are not wealthy. The relatively low progressivity tax for labor favors the high-wage and thus high-productivity agents, which implies bigger weights for these agents, however, for agents who are credit constrained the presence of the progressivity favors them and does not cause distortions, which implies the planner gives a higher weight for those agents.

For France, the shape of Pareto weights is different from the US one. Although it has also the U-shape format, one can note that it favors more the low productivity agents. The weights are also sharply decreasing with productivity and involves less heterogeneity in Pareto weights than the US distribution. This shape for Pareto weights in the income dimension result from a higher progressivity for the labor tax. As it was the case with the US, the heterogeneity along the wealth dimension is much more limited, although we can see it varies less than in the US case. Notice that for high levels of productivity/income, the weights along the wealth dimension become less heterogeneous. The same can be said when we analyze the evolution of the weights for higher levels of wealth, where we can notice the weights become more homogeneous as we increase the productivity/income. This result corroborates to the idea that the fiscal system in France shows a higher inequality aversion component than the one in the US.

---

where the values of $\frac{K_t}{Y_t}$ and $\frac{B_t}{Y_t}$ were calculated previously. For both countries we use the GDP data measured at national currency, current prices. The data for the GDP was obtained in the OECD for the year 2007.



(a) United States (b) France

Figure 10: Pareto weights as a function of history for the US and France.

Finally in Figure 10 we plot the Pareto weights for the US in panel (a) and France in panel (b) as a function of histories. The blue dashed line represent those weights calculated using equation (C.56) in Appendix C. In this case we have as many Pareto weights as the possible number of histories. In the solid red line we present the same weights using equation (C.55) in Appendix C. Those weights are calculated such that histories with the same productivity in the beginning of the truncation are given the same weight. In this case we will be let with 10 possible Pareto weights. The histories are organized in increasing order of the productivity level in the beginning of the truncation, which means the red solid line represents the weights showed previously in Figure 2. One can notice that although there are differences between those two ways to compute the weights, they are globally very similar.

E Consistency of the results

E.1 $N = 7$

First we consider the case where the number of truncation histories is $N = 7$. Figure 11 shows the same result as in Figure 9.

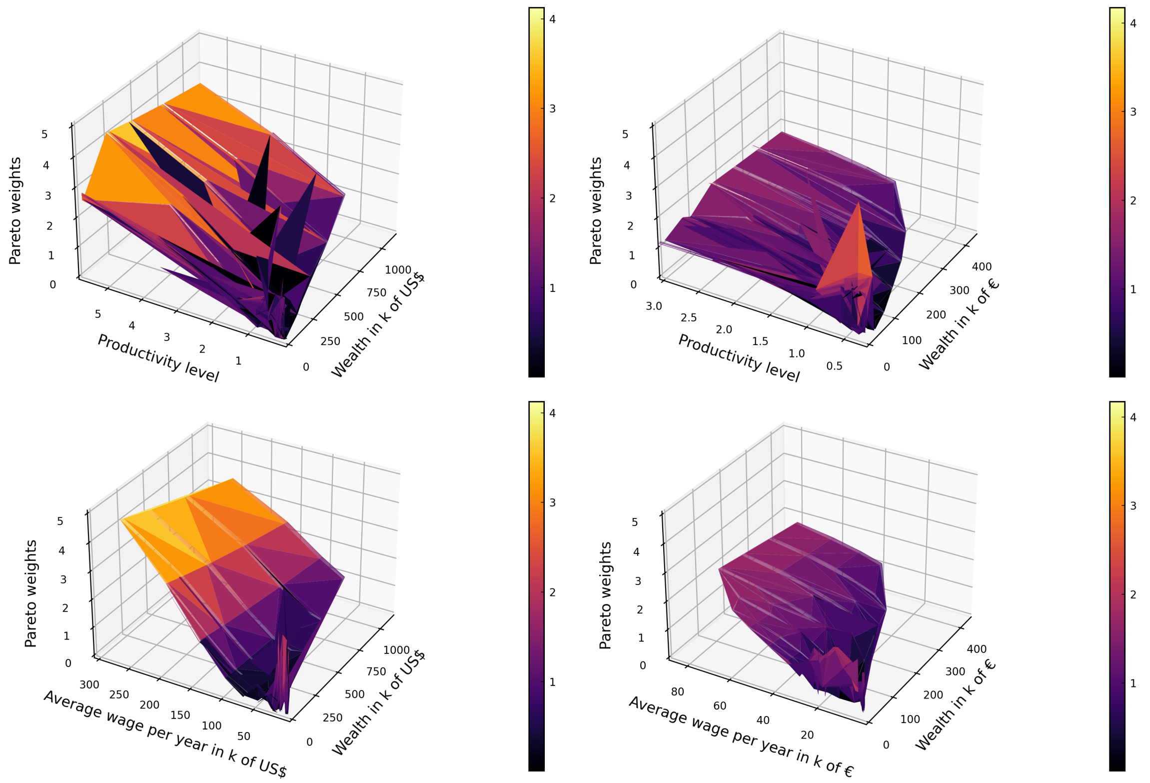
As we did in Figure 2, the Pareto weights along the productivity dimension is plotted in Figure 12.

We also plot the results along the wealth dimension in Figure 13.

Table 9 shows some summary statistics for the Pareto weights in this scenario.

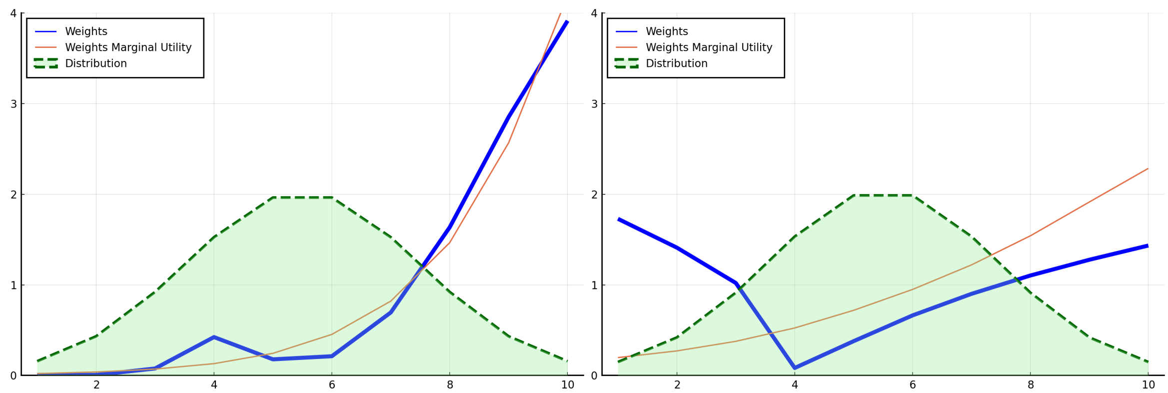
	USA	**France**
Mean	1.00	1.00
St. deviation	1.37	0.508
Min.	0.006	0.0832
Max.	3.91	1.73
Bottom 10 %	0.006	0.35
Median	0.32	1.06
Top 10%	2.96	1.46

Table 9: Summary statistics for the Pareto Weights of the US and France N=7.



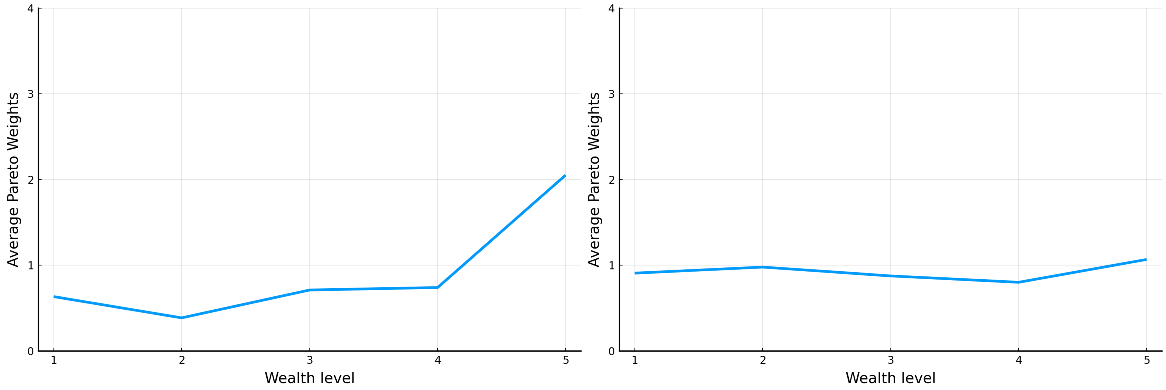
(a) United States (b) France

Figure 11: Pareto weights as a function of productivity (income per capita) and wealth for the US and France N=7.



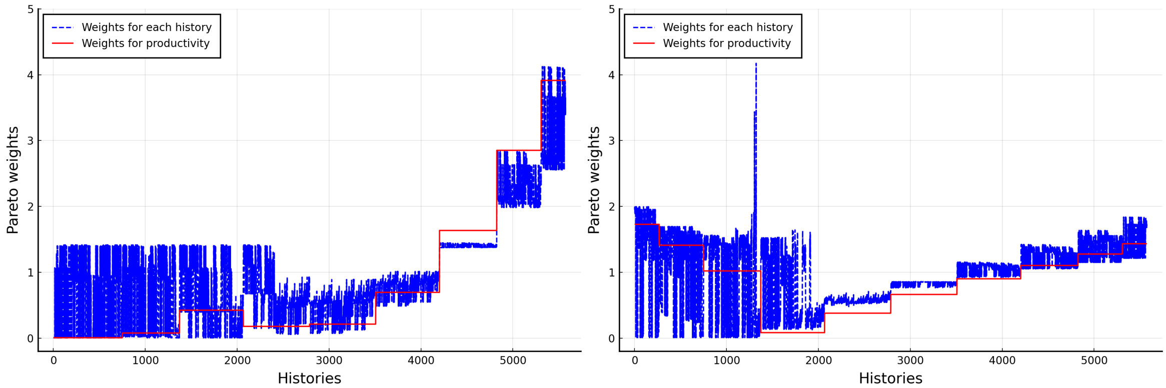
(a) United States (b) France

Figure 12: Pareto weights as a function of productivity for the US and France N=7.



(a) United States (b) France

Figure 13: Average Pareto weights as a function of wealth for the US and France N=7.



(a) United States (b) France

Figure 14: Pareto weights as a function of history for the US and France N=7.

Lastly as in Figure 10 we plot in Figure 14 the Pareto weights for the US in panel (a) and France in panel (b) as a function of histories for truncation length of $N = 7$.

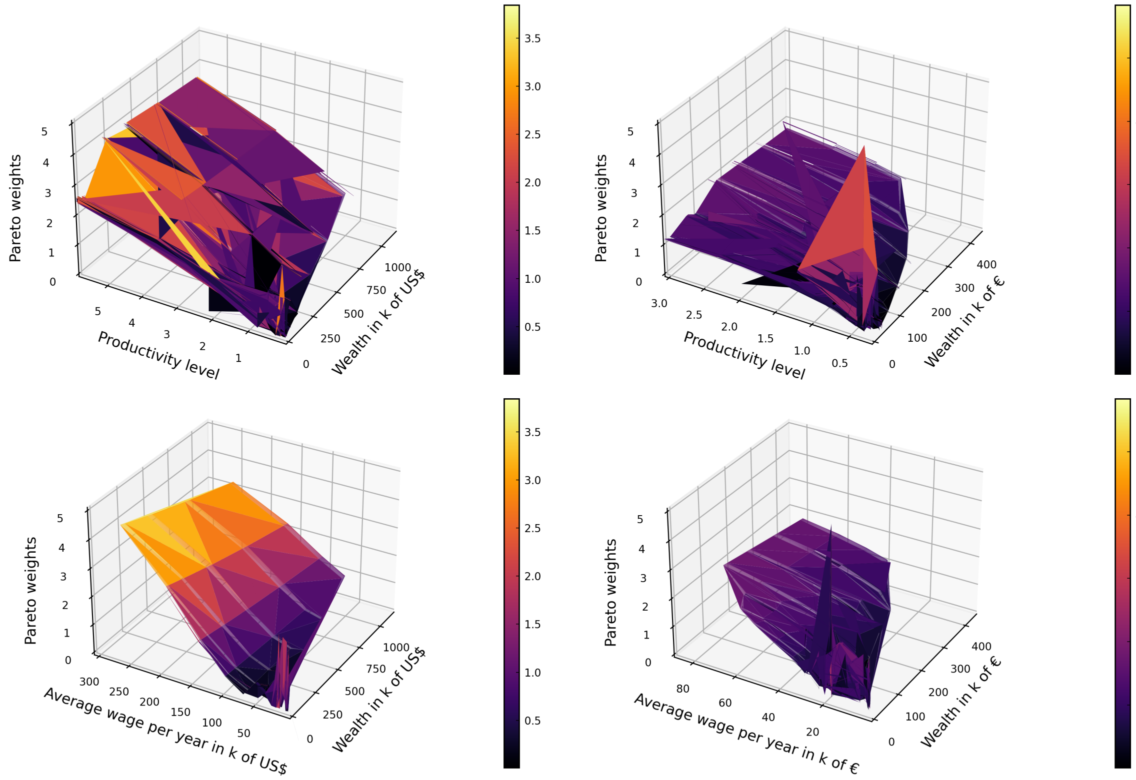
E.2 $N = 8$

Now we consider the case where the number of truncation histories is $N = 8$. Figure 15 shows the same result as in Figure 9.

As we did in Figure 2, the Pareto weights along the productivity dimension is plotted in Figure 16.

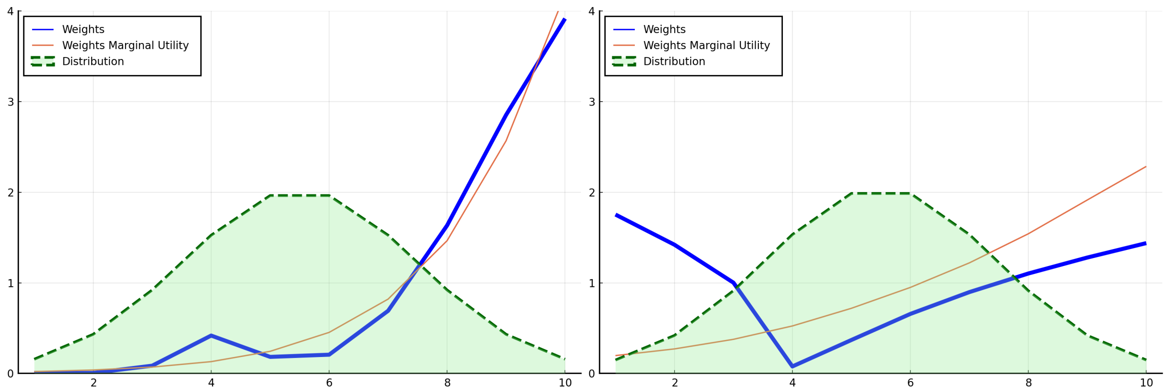
We also plot the results along the wealth dimension in Figure 17.

Table 10 shows some summary statistics for the Pareto weights in this scenario.



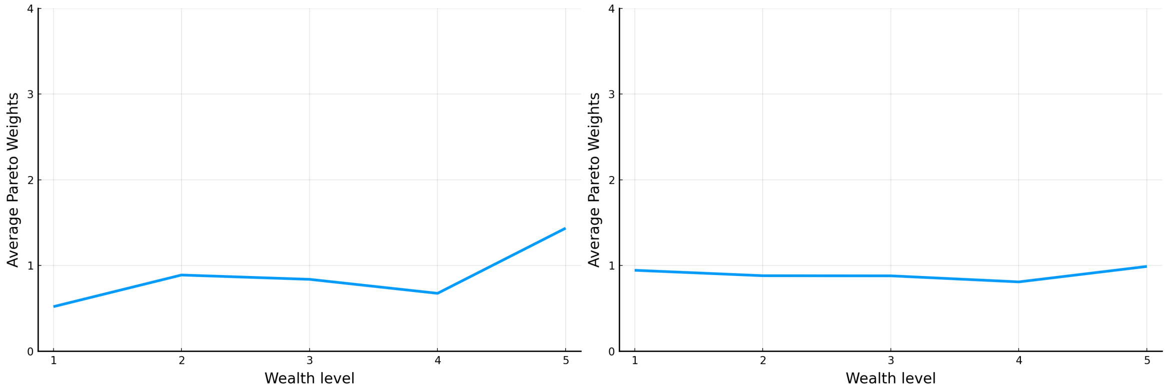
(a) United States (b) France

Figure 15: Pareto weights as a function of productivity (income per capita) and wealth for the US and France N=8.



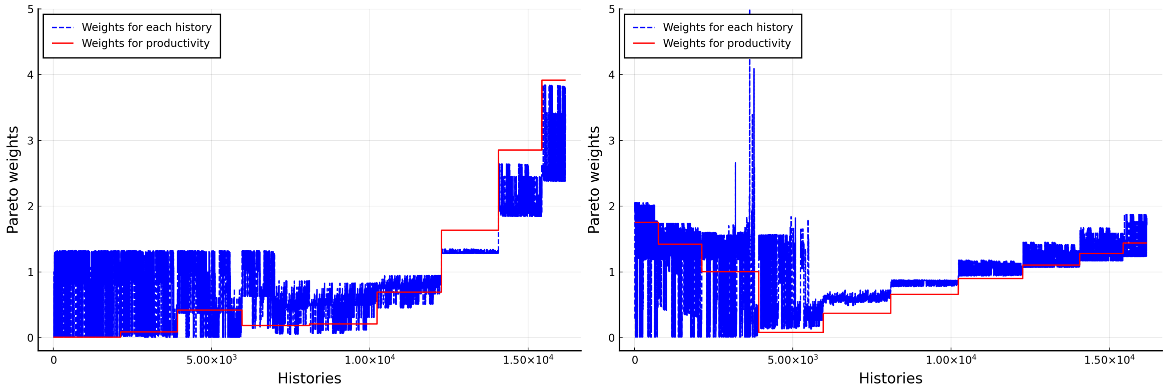
(a) United States (b) France

Figure 16: Pareto weights as a function of productivity for the US and France N=8.



(a) United States (b) France

Figure 17: Average Pareto weights as a function of wealth for the US and France N=8.



(a) United States (b) France

Figure 18: Pareto weights as a function of history for the US and France N=8.

	USA	**France**
Mean	1.00	1.00
St. deviation	1.37	0.517
Min.	0.006	0.0779
Max.	3.92	1.75
Bottom 10 %	0.006	0.34
Median	0.31	1.05
Top 10%	2.96	1.47

Table 10: Summary statistics for the Pareto Weights of the US and France N=8.

Lastly as in Figure 10 we plot in Figure 18 the Pareto weights for the US in panel (a) and France in panel (b) as a function of histories for truncation length of $N = 8$.

F Changes in Pareto weights by increasing G/Y

In this Appendix we test the robustness of our estimates through a simple exercise. We increase the government spending to output ratio of United States by 10% and of France in 1%. We then conduct different experiments financing this increase by each one of the instruments (i.e., τ_k, τ_c, κ, and τ) such that the budget constraint of the state is still satisfied. For each one of the instruments we plot the difference in the weights using our estimation strategy discussed previously. This result is showed in Figure 5.

Table 6 shows the changes in the instruments such that the budget of the state is satisfied after an increase in G/Y. Notice that for each one of the cases we need to increase the tax on capital (τ_k), tax on consumption (τ_c), increase the average tax on labor ($1-\kappa$), and increase the progressivity (τ). It is worth noting that finance the increase in G by either τ_c or κ has the same effect in terms of general equilibrium, which implies the difference in weights is the same for both experiments. This result is linked to the redundancy result we discuss in Section 5.1. Finally one can observe that the instrument that leads to the highest change in the weights is the progressivity of the labor tax and this result corroborates to the idea that by increasing progressivity we benefit agents with low productivity shocks at the expense of those at the top, although even in this case the Pareto weights keep the same shape as the ones showed in Figure 2.

Figure 19 illustrates the changes in Pareto weights considering its parametric version, $\log \omega(y) := \omega_0 + \omega_1 \log(y) + \omega_2(\log(y))^2$, as well as the changes in the utility of the agents, labor, and capital income for each experiment explained above. One can notice that as expected the increase in tax on capital leads to a reduction of the capital income, especially for agents with high productivity. This reduction on capital income encourages them to increase the amount of hours worked, which leads to an increase in the labor income. However, this effect is not enough to compensate the lost of revenues arising from the increase in the tax on capital and so the utility for these agents is reduced.

The increase in the tax on consumption has a high effect in the labor income, which is reduced. Notice that the capital income does not change much for all the agents. The reduction in utility for the low productivity agents is higher in this case, probably due to the fact that these agents are credit constrained and cannot compensate the lost in purchase power by relying on assets. The κ has the same effect of tax on consumption. For all the cases analysed so far the changes in the utility of the agents is not so drastic and as a result the planner does not alter the weights in a substantial way.

Finally notice that an increase in progressivity, increases the labor income for the agents with low values of idiosyncratic states and reduce of those with high values. In the latter case, the agents now have less incentives to save, which reduces their capital income. The overall effect is then an increase in the utility of agents at the bottom of income distribution and a decrease for those at the top. This effect is captured by the planner, which increases the weights of the low productivity agents at the expense of the high productivity ones. Figure 20 shows the same case for France and the analysis is similar to the one discussed above for the US case.

Figure 21 shows the change in the Pareto weights for the United States for the experiment we increase the value of G/Y. The blue line represents the Pareto weights in the steady state and the red line the experiment conducted above, where the panel (a) designates the Pareto weights we obtain after an increase in G financed by tax on capital, panel (b) by tax on consumption, panel (c) by the average tax on labor, and panel (d) by progressivity. Notice that although, the case where we finance the increase in G by increasing the progressivity shows the highest difference in terms of the weights, the shape is the same. Figure 22 is this analysis for France.

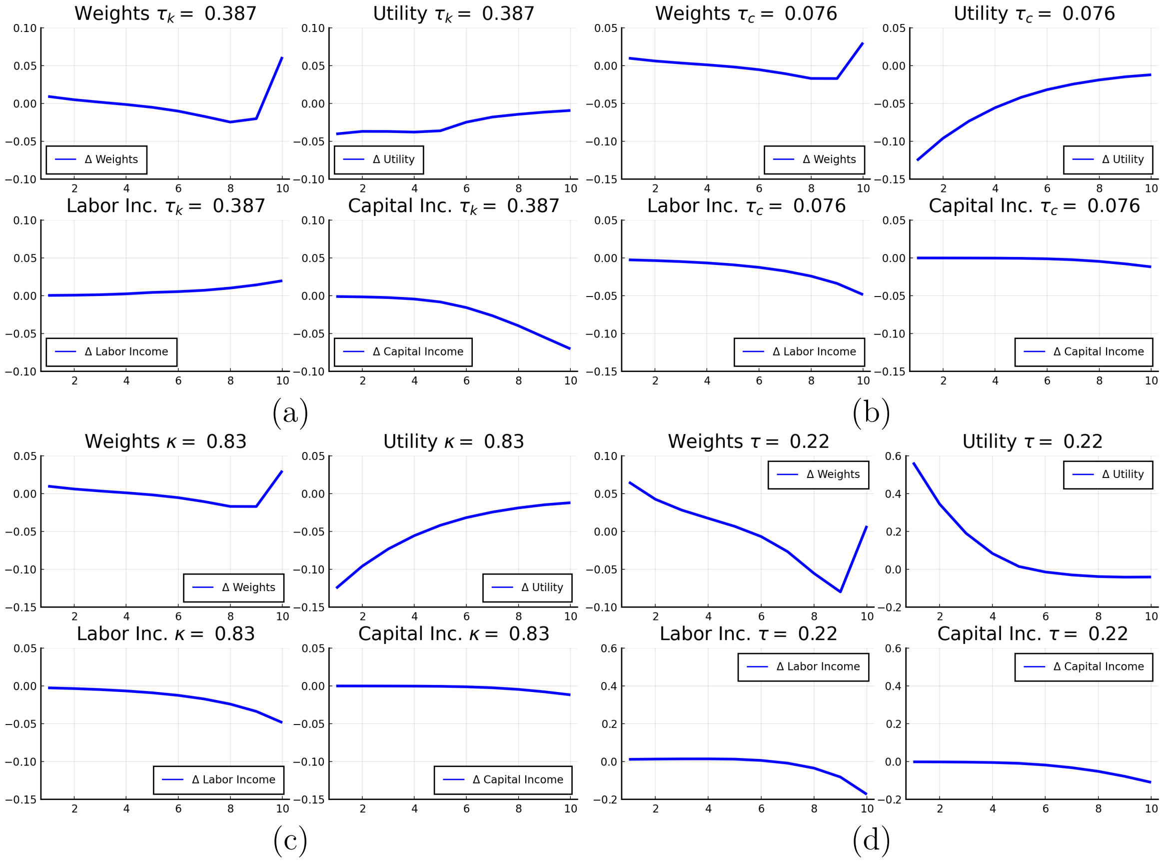


Figure 19: Difference in parametric Pareto weights, utility, labor, and capital income after an increase in G/Y financed by different instruments for United States.

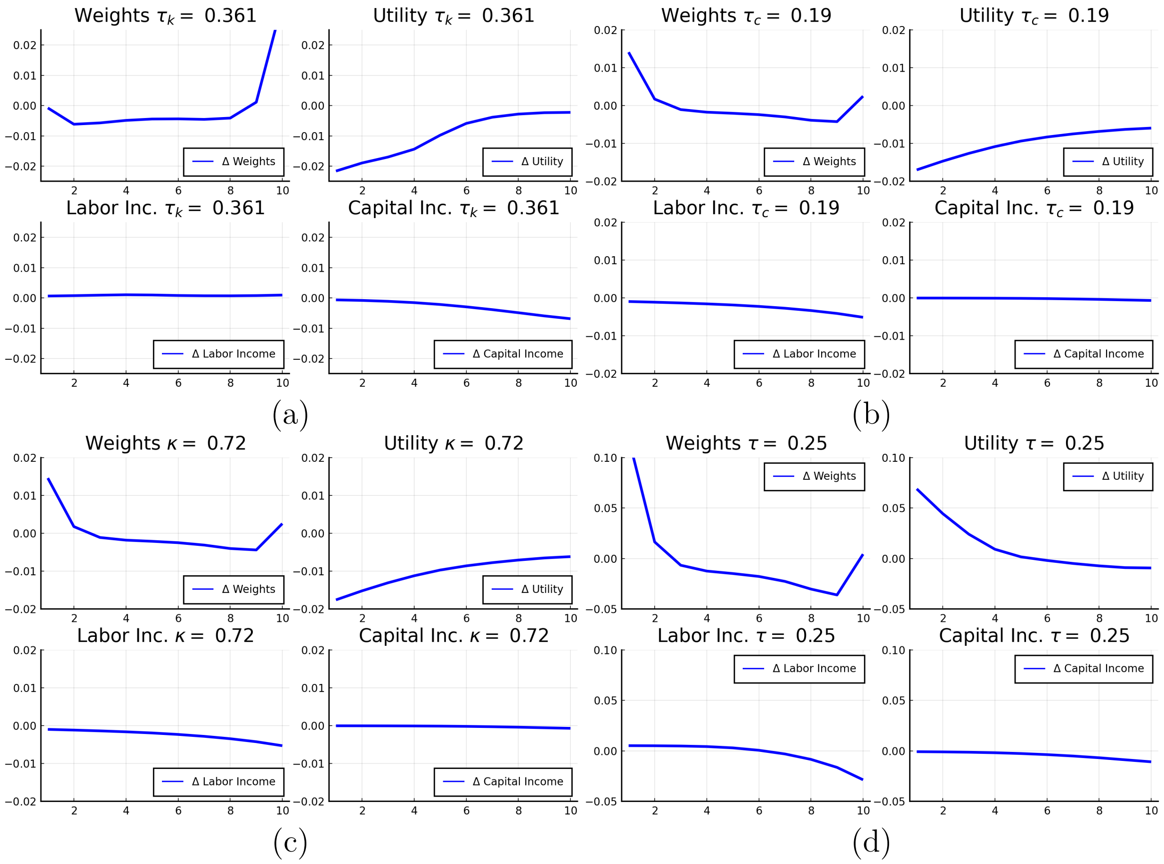


Figure 20: Difference in parametric Pareto weights, utility, labor, and capital income after an increase in G/Y financed by different instruments for United States.

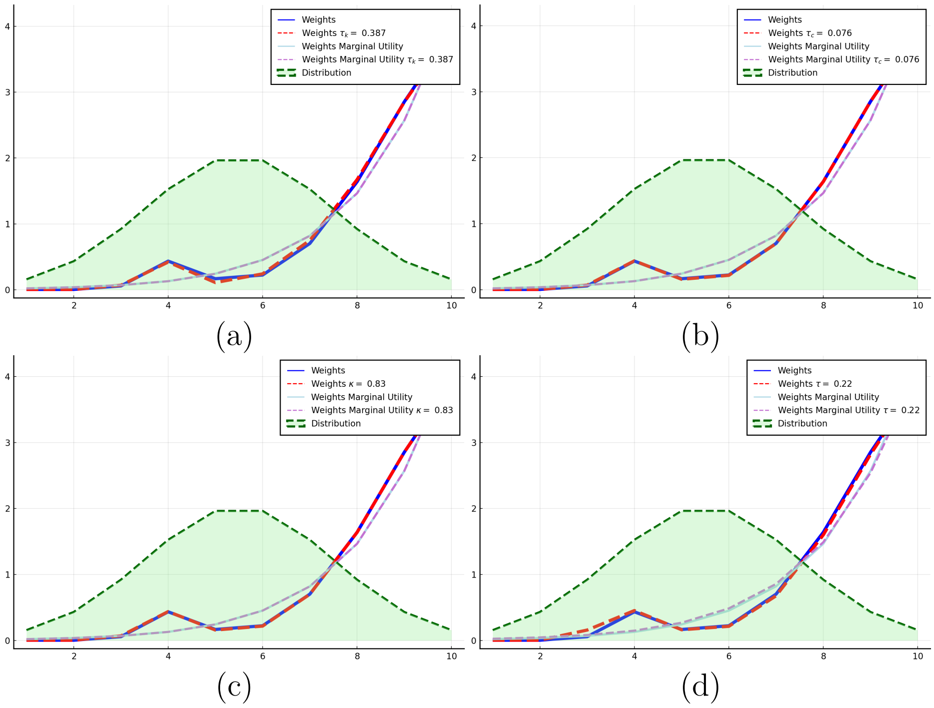


Figure 21: Pareto weights for United States after and increase in G/Y being financed by different instruments.

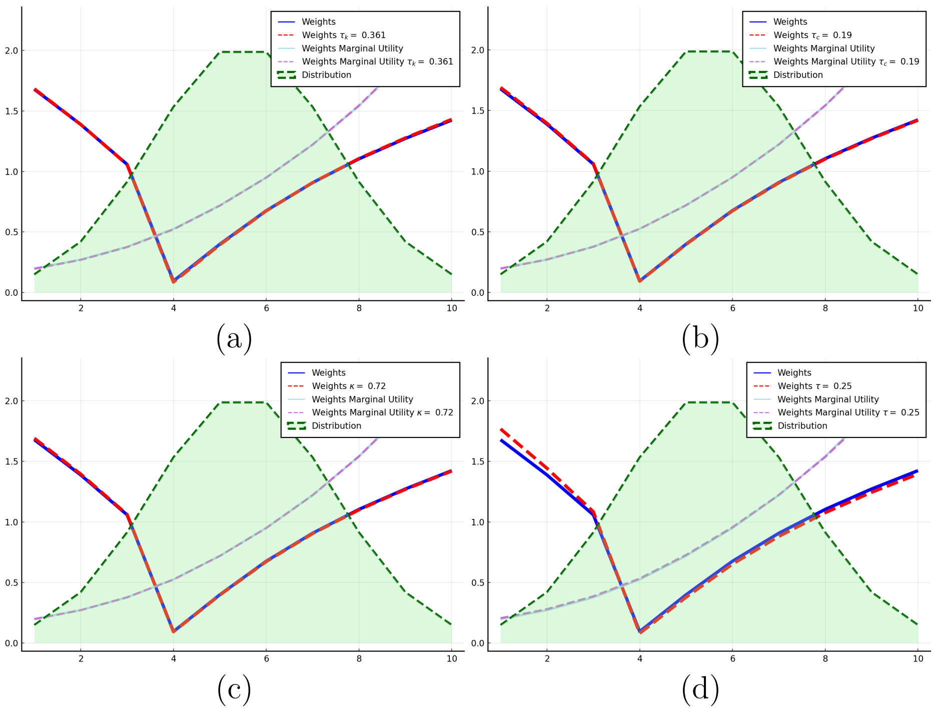
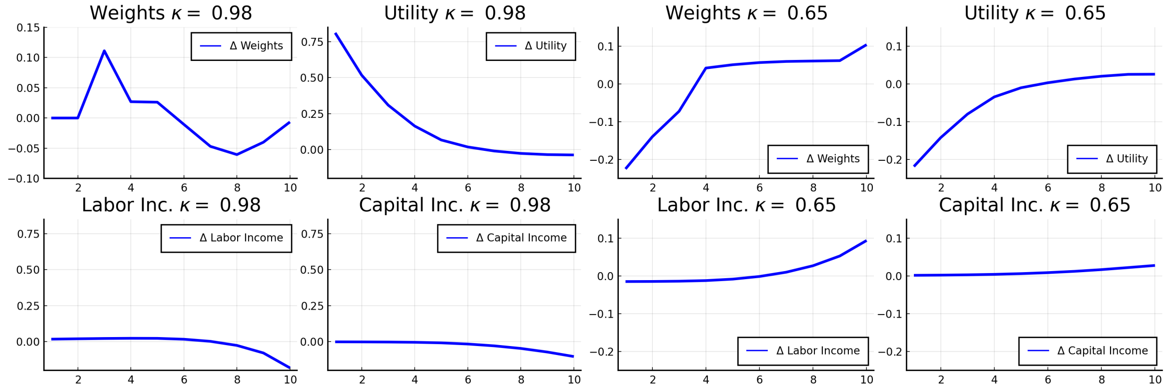


Figure 22: Pareto weights for France after and increase in G/Y being financed by different instruments.



(a) United States (b) France

Figure 23: Difference between the Pareto weights between US with the French fiscal system.

G Changes in the Fiscal system

Figure 23 shows how the Pareto weights, utility, labor, and capital income change when we alter the fiscal system of each country by the other one under consideration. In panel (a) we show the case where United States adopt the tax system of France and in panel (b) is the case when France adopts the tax system of United States. As expected we can notice the effects occur in opposite directions. The experiment showed in Figure 23 is conducted as follows, once the capital-to-output ratio is set to the value in the steady state, we iterate in the value of κ such that G/Y is kept the same. By running the experiment in this way not only the model parameters are unchanged but also the main macro ratios.

Table 11 shows the new fiscal system for each country under consideration.

	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 11: Fiscal system for United States and France under the experiment.

H Evolution of the main variables in the experiment

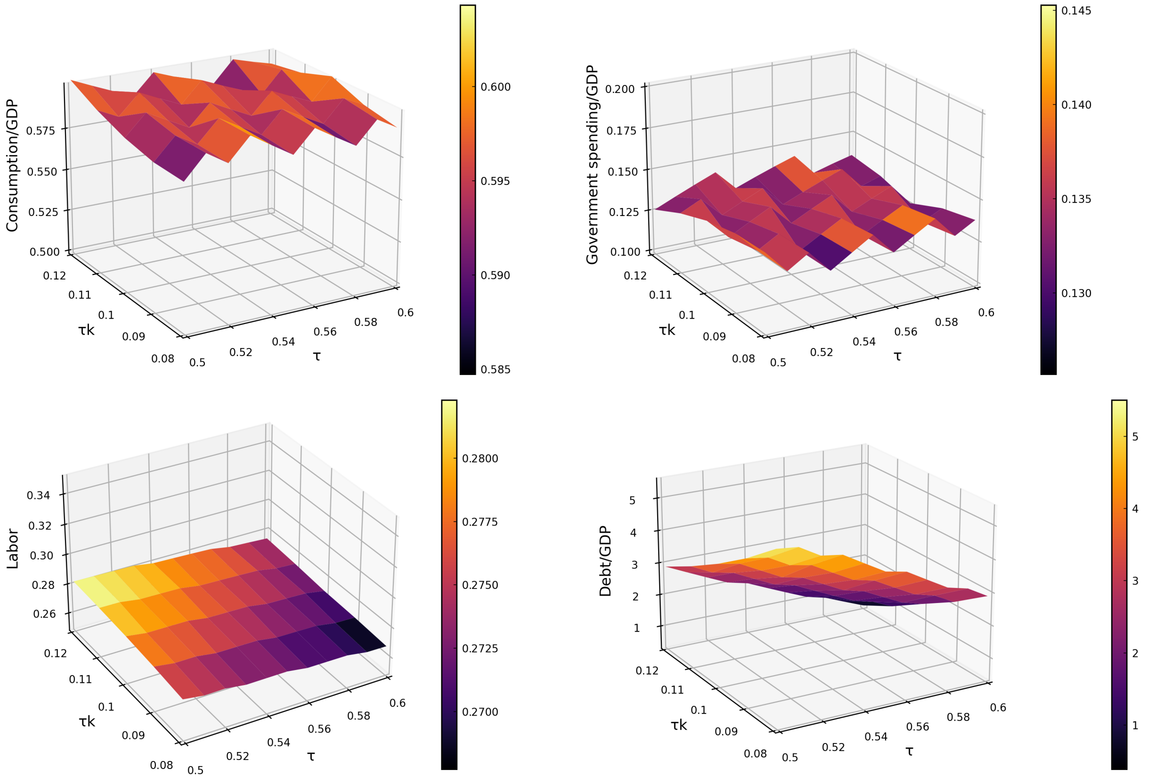


Figure 24: Evolution of the main variables along the τ_K and τ dimension.