The Evolution of the Welfare State

Diego Huerta [*]

November 9, 2023

[Latest version]

Abstract

The evolution of the welfare state over the last few decades has differed strikingly across countries in the world. For example, spending on social benefits as a fraction of GDP has substantially increased in the US since 1980, remained stable in Canada, and declined in Sweden. To explain these different trends, I propose a model with agents that are heterogeneous in occupation and wealth, and who vote on social benefits over the course of their lifetime. The model highlights the key role of "aspirational voters"—members of the middle class who support pro-business policies and sacrifice social benefits hoping to become future entrepreneurs. The importance of aspirational voters, in turn, depends on wealth inequality. The model predicts that social spending should increase in rich countries with high wealth inequality, while it should decline if inequality is low. A calibrated version of the model successfully predicts the observed trends of social spending in 18 out of 24 countries from all continents.

[*]Department of Economics, Northwestern University, 2211 Campus Drive, Evanston, Illinois 60208, USA. Email: diegohuerta2024@u.northwestern.edu. Web: diegohuertad.com.

I would like to thank to Giorgio Primiceri, Georgy Egorov, and Larry Christiano for their invaluable guidance and support. I also thank Alessandro Pavan, Nicola Persico, Marios Angeletos, Kiminori Matsuyama, Matthias Doepke, Bruno Strulovici, Ronald Fischer, Saki Bigio, Timothy Feddersen, Daniel Barron, Benjamin Friedrich, Meghan Busse, Elisa Jacome, Ryan Hill, Lulu Wang, Walker Hanlon, Matias Bayas-Erazo, Amilcar Velez, Fergal Hanks, Joao Guerreiro, Alan Liu, and Nicolas Garrido for very useful comments and discussions. I also thank the participants of the NU Macro lunch and of the Brown Bag Seminar at Kellogg for helpful comments.

1 Introduction

International data on spending on social benefits from the OECD reveals striking differences in the evolution of the welfare state across countries in the last few decades. To illustrate this phenomenon, Figure 1 presents the dynamics of social benefits as a fraction of GDP in three countries: the United States, Canada, and Sweden. The United States has witnessed a persistent increase in social spending since 1980, Canada has maintained relatively stable levels, but the trend has been decreasing in Sweden since 1995. These three countries are representative of the kind of patterns observed across the world. For instance, social benefits have also increased over time in Australia, remained stable in Ireland, and decreased in Israel.

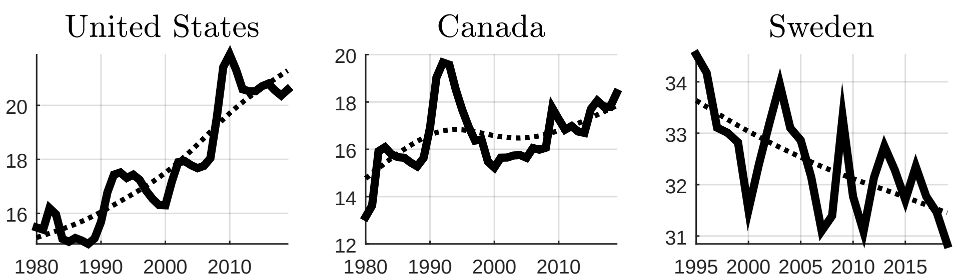


Figure 1: Solid line: spending on social benefits (% of GDP). Dotted line: trend. Source: OECD.

These diverging trends in social benefits can be due to many possible factors such as country-specific shocks, changes in government administrations, demographics, or technological progress. In this paper, I focus on the role of a single driver: the wealth distribution.

The notion that the wealth distribution affects the size of the welfare state has a long history in the social sciences. De Tocqueville (1835), for example, stressed that the relatively high wealth inequality in Europe compared to the United States increased class conflict, creating the need for a large government to reduce inequality. In this paper, I propose a tractable model that captures the centrality of the wealth inequality for the evolution of the welfare state. Despite its parsimonious nature, my model successfully predicts the recent trends of social spending in 18 out of 24 countries from all continents, including the three countries presented in Figure 1, and also other countries such as Australia, Israel, and South Africa.

My model works as follows. There is no uncertainty. The economy is populated by a continuum of agents who are heterogeneous in wealth and occupation. In each period, they choose whether to be *workers* who supply labor, or *entrepreneurs* who operate a firm that produces physical capital subject to collateral constraints. A representative firm manufactures the final product by hiring workers and purchasing the physical capital from the entrepreneurs.

In each period, agents vote for candidate governments that propose a transfer package to maximize their share of votes. Candidates face a crucial trade-off between providing transfers to workers or entrepreneurs. The size of the welfare state corresponds to the share of GDP that is used to finance transfers to workers. More social benefits to workers require lower transfers to entrepreneurs, tightening collateral constraints, and thus, decreasing entrepreneurship (for recent evidence see Audretsch et al., 2022; Solomon et al., 2022). Transfers are financed through a wealth tax to all agents while maintaining a balanced budget.

In each period, citizens make three sequential decisions. First, they vote on government transfers while anticipating the effects of policies on their occupational decisions. Second, based on their wealth and on the current size of the welfare state, agents make their occupational choice. Finally, individuals decide on consumption and savings. These individual saving decisions, in turn, determine the future wealth distribution, shaping the outcome of future voting rounds.

In their decision-making process, agents expect social benefits to remain stable in the future due to political constraints, inefficiencies in the tax system, or difficulties in the allocation of public funds. While this assumption deviates from the common notion of fully rational agents, it allows for an analytical characterization of transition dynamics, which is the main focus of the paper. This approach is akin to the one used in anticipatory utility models in the learning literature (Sargent, 1999) and similar to the temporary equilibrium concept employed in housing models (Piazzesi and Schneider, 2016). Without this assumption, the model becomes highly untractable. This is the reason why the papers studying the fully rational equilibrium have so far relied on numerical solutions (Krusell and Rios-Rull, 1999).

To understand how the wealth distribution shapes the voting outcome, it is useful to study the individual preferences for social benefits as a function of wealth. I classify agents into three classes according to their preferences and occupational prospects: the working, the emerging, and the incumbent class. The working class represents people at the bottom of the wealth distribution. They find it optimal to demand high social benefits, anticipating that their resources are insufficient to overcome collateral constraints and start a firm. The emerging class includes middle wealth agents who are willing to sacrifice social benefits and support pro-business policies aspiring to become entrepreneurs (referred to as "aspirational voters"). The incumbent class consists of high wealth individuals. These agents prefer less pro-business policies and may be even willing to pay for social benefits to limit the competition from the emerging class who wish to enter the market.

These individual preferences are aggregated through a probabilistic voting model à la Persson and Tabellini (2000). The equilibrium size of the welfare state maximizes the current politically-weighted income of workers and entrepreneurs. The "political weight" can be interpreted as capturing the political orientation of a representative government, either more pro-worker or pro-business. The equilibrium level of social benefits depends crucially on the "strength" of aspirational voting, that is, on the mass and intensity of the preferences of the emerging class. Thus, the evolution of the size of the welfare state hinges on the dynamics of aspirational voting.

To tackle the challenges associated with characterizing transition dynamics in models with heterogeneous agents, I restrict attention to a constrained set of initial distributions. My main theoretical result is that the evolution of the welfare state depends on a country's initial aggregate wealth and inequality. The model predicts that social spending should increase in rich countries with high wealth inequality, while it should decline if inequality is low. These predictions are in line with the American and Swedish experiences illustrated in Figure 1.

To grasp the main intuition behind these results, consider a country that starts rich but unequal. High inequality implies that few agents have sufficient wealth to start a firm. Thus, markets are not very competitive and firms' profits are high, making entrepreneurship very attractive. As a consequence, there is an initially large mass of aspirational voters who support low social spending. As the economy accumulates wealth, more agents are able to start a firm, causing profits to decline. As a result, aspirational voting weakens over time, leading to an increasing path of social benefits.

I test the predictions of the model on data since 1995 from 24 countries from all continents. I calibrate the model parameters to match the level of social benefits in each country in 1995, taking as given the observed wealth distribution in that same year. Then, I simulate the model for the subsequent 25 years to assess its ability to predict the observed trends of social benefits. The only exogenous force hitting the economy is the aggregate productivity in the production function of the representative firm, which is estimated via Solow residuals for each country.

The model correctly predicts the sign of the trend of social benefits in 18 out of the 24 countries, and in many cases also the magnitude and shape of the trend. These experiments provide strong empirical support to the conclusion that the wealth distribution is a key force behind the striking differences in the evolution of the welfare state across countries.

To assess the importance of changes in the political orientation of the government for the evolution of the welfare state, I perform a counterfactual analysis in Canada, the United States, and Sweden . For each country, I solve for the path of political weights that matches the observed path of social benefits. Then, I simulate a "highly pro-worker" and a "highly pro-business" scenario. I find that the trend of social benefits would not have changed significantly under both scenarios. This confirms the key role of the wealth distribution in explaining the dynamics of the welfare state.

1.1 Literature Review

This paper adds to a vast literature on endogenous policy choice and redistribution, starting with the seminal work of Meltzer and Richard (1981). In general, high inequality has been associated with more extensive welfare states (e.g. Alesina and Rodrik, 1994; Persson and Tabellini, 1994). Alesina and Angeletos (2005) and Hassler et al. (2005, 2003a,b) study the long-run survival of the welfare state. My paper contributes to this literature in at least two ways. First, it shows that countries with high inequality can, in fact, have low social benefits due to the key role of aspirational voters.[1] Second, to my knowledge, is the first paper to test a political theory for the welfare state through a quantitative analysis in a set of countries, demonstrating its ability to predict the observed trends.

This article also contributes to a macro literature that introduces politics to heterogeneous agents models, pioneered by Krusell et al. (1996).[2] These papers rely on numerical methods to study a fully rational equilibrium where agents anticipate the future evolution of all variables. I add to this literature by proposing a tractable model with heterogeneous agents, occupational choice, and politics that allows for a theoretical characterization of the transition dynamics.[3]

This paper also contributes to a recent literature that provides theoretical results for optimal policy interventions in models with heterogeneous agents (Acharya et al., 2023; Itskhoki and Moll, 2019; Nuño and Moll, 2018).[4] In particular, Itskhoki and Moll (2019) show that the optimal Ramsey policy starts pro-business at the early stages of development, and switches to pro-worker in the long-run. My model shows that such policy dynamics can also arise in equilibrium when policies are chosen through repeated voting over time. My model also provides a political rationale for countries that do not fit the normative theory, such as Sweden in the last three decades.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 defines the equilibrium. Section 4 studies the transition dynamics when the size of the welfare state is exogenous and fixed over time. Section 5 presents the political process that defines the equilibrium size of the welfare state. Section 6 characterizes the evolution of the welfare state over time. Section 7 presents the quantitative results. Section 8 concludes.

[1]Related to my result of a middle-class that supports pro-business policies (aspirational voters), Benabou and Ok (2001) examine the prospect of upward mobility hypothesis (POUM) to explain why the poor may not support high redistribution. Related work also includes Acemoglu et al. (2018); Leventoğlu (2014); Benabou and Tirole (2006); Alesina and La Ferrara (2005); Leventoğlu (2005); Piketty (1995).

[2]See also Krusell and Rios-Rull (1999); Krusell et al. (1997). More recent literature includes: Quadrini and Rios-Rull (2023); Jang (2023); Pecoraro (2017); Bachmann and Bai (2013)

[3]Papers that work quantitatively with heterogeneous agent models and occupational choice but that do not incorporate a political dimension include: Buera and Shin (2013); Buera et al. (2011); Cagetti and De Nardi (2006). My model also relates to the seminal work by Banerjee and Newman (1993) and Galor and Zeira (1993).

[4]Recent literature that obtains analytical results in heterogeneous agent models but whose focus is not optimal policies include Achdou et al. (2022); Buera and Moll (2015); Moll (2014). This paper also relates to a literature that studies Ramsey taxation under self-interested politicians (Acemoglu et al., 2011, 2010, 2008; Yared, 2010).

2 Model

2.1 Preferences

Time is continuous, there is an infinite time horizon, and no uncertainty. The economy is populated by a continuum of individuals who are heterogeneous in wealth, a. At each point in time, all agents are endowed with ℓ units of labor. The state of the economy at period t is given by the wealth distribution function, $\gamma_t(a)$. The cumulative wealth distribution is denoted by $\Gamma_t(a)$. Agents have standard preferences over utility flows from consumption c_t with a discount rate $\rho \geq 0$:

$$\mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \log(c_t) dt. \tag{2.1}$$

The assumption of a logarithmic utility simplifies the theoretical characterization of transitions dynamics, but can be generalized to a constant risk aversion utility. Section B.2 in the Appendix presents this extension.

2.2 Technology

There are different production technologies for output and physical capital (e.g Bernanke and Gertler, 1989; Matsuyama, 2004). A representative firm produces the single output good according to a Cobb-Douglas production function: $F(K_t, L_t) = ZK_t^{\alpha} L_t^{1-\alpha}$, with $\alpha \in (0,1)$. K_t and L_t are aggregate physical capital and the aggregate labor supply, respectively. The factor markets are competitive. Z is aggregate productivity which can potentially follow an exogenous time path.[5]

In each period, individuals make an occupational choice. Agents have two options. First, they can become workers and supply their labor ℓ to the representative firm. In that case, they receive a labor income given by: $w_t \cdot \ell$, where $w_t = (1-\alpha)ZK_t^{\alpha} L_t^{-\alpha}$ is the wage rate.

Second, they can decide to invest in a capital-producing firm and become entrepreneurs. The capital production technology requires $I > 0$ units of investment and ℓ units of labor to produce

[5] In the theoretical analysis, I assume that Z is fixed over time. In Section 7, where I present the quantitative exercise, I allow the aggregate productivity to follow an exogenous path.

$R > 0$ units of physical capital. Entrepreneurs use their own labor ℓ for the production of capital. There is no physical capital accumulation technology.[6] Thus, all capital produced is used for the production of the output good. In particular, entrepreneurs sell their physical capital to the representative firm at a price $p_t = \alpha Z K_t^{\alpha-1} L_t^{1-\alpha}$. Entrepreneurial profits are given by: $\Pi_t = p_t \cdot R - rI$.

Throughout the paper, I denote by e_t the fraction of agents that become entrepreneurs at period t. Then, total physical capital is: $K_t = R \cdot e_t$. Aggregate labor supply is: $L_t = \ell \cdot (1 - e_t)$. Total output is given by: $Y_t = F(Re_t, \ell(1 - e_t))$.

2.3 The Size of the Welfare State

The size of the welfare state corresponds to the fraction b_t of total income Y_t that is used to finance transfers to workers, i.e. spending on social benefits.[7] Thus, each worker receives a transfer given by: $T_t = b_t \cdot Y_t$. Throughout the paper, I refer to b_t as the "transfer rate". The equilibrium size of the welfare state is decided through repeated voting over time. In particular, agents vote in each period for two candidate governments that propose simultaneously a transfer rate to maximize their share of votes. Section 5 describes the political process that takes place in each period.

Social benefits are financed by levying a wealth tax, τ_t, to all agents while keeping a balanced budget:[8]

$$\tau_t \cdot A_t = T_t \cdot (1 - e_t), \tag{2.2}$$

where A_t is the aggregate wealth at period t: $A_t = \int a d\Gamma_t(a)$, while $(1-e_t)$ is the fraction of agents that become workers. Hence, condition (2.2) equalizes total tax revenues and total transfers used to finance social benefits. Section B.4 in the Appendix shows that, under some conditions, the main results remain valid when social benefits are financed through both a labor and a wealth tax.

I assume that the transfer rate is bounded from below by $-\underline{b}$, i.e. $b_t \geq -\underline{b}$. I impose the

[6] Alternatively, physical capital fully depreciates after one period.

[7] A similar modelling approach has been used in the political economy literature on the welfare state (e.g. Hassler et al., 2003a).

[8] The tax rate can be also interpreted as a capital income tax.

following assumption on the minimum transfer rate:

Assumption 1

$$-\underline{b} \geq -\frac{rI}{ZR^{\alpha}\ell^{1-\alpha}}.$$

This is a sufficient condition to provide a theoretical characterization for the evolution of the equilibrium transfer rate, b_t. In most cases, this lower bound is not restrictive, i.e. $b_t > -\underline{b}$. However, it is required in some particular cases to guarantee convergence to a steady state transfer rate (see Section 6). When the transfer rate is negative, entrepreneurs receive net positive transfers that are financed by workers (referred as to "business policies").[9]

Overall, a candidate government faces a crucial trade-off: providing social benefits to workers or spending on business policies that benefit entrepreneurs. In Section D.2, I present preliminary evidence for the European Union that supports the negative relationship between spending on social benefits and business policies over time. Additionally, in Section B.3 in the Appendix, I study a more realistic policy instrument in which the government can decide to provide positive transfers to both workers and entrepreneurs. The theoretical results of the paper still hold under some additional assumptions on the exogenous parameters of the model.

2.4 Budgets

Workers Agents take the wage rate w_t as given and can borrow and save at the fixed international interest rate: $r_t = r$. The individual wealth for agents that decide to become workers evolves according to:

$$\dot{a}_t = (r - \tau_t)a_t + w_t\ell + T_t - c_t. \tag{2.3}$$

[9]Note that if $b_t < 0$, then $T_t < 0$ and $\tau_t < 0$. Thus, in that case entrepreneurs receive a capital subsidy, while workers bear the cost. In particular, I assume that the lower bound on b_t is relatively low (e.g. $\underline{b} = -100\%$).

Entrepreneurs Agents that decide to start a capital-producing firm take the price of capital p_t as given and receive profits: $\Pi_t = p_t \cdot R - rI$. Thus, they face the flow budget constraint

$$\dot{a}_t = (r - \tau_t)a_t + \Pi_t - c_t. \tag{2.4}$$

All individuals, workers and entrepreneurs, also face a borrowing limit $a_t \geq \underline{a}$, where $-\infty < \underline{a} \leq 0$. This borrowing limit guarantees the convergence to a steady state distribution in some cases (see Section 6).

2.5 Credit conditions

In order to finance a firm, entrepreneurs may apply for a loan, $d_t = I - a_t$. There is a zero-profits banking system which provides loans and has unlimited access to international funds at the fixed interest rate r.

There is a moral hazard problem in the credit market: agents may abscond with the loan and become workers instead of investing in a firm. To prevent such malicious behavior, banks ask agents to deposit their wealth at the beginning of each period. If the agent defaults, as a punishment, she would lose her deposited collateral, but she would receive the social benefits she is entitled as a worker.

In order to receive a loan, an agent with assets a must satisfy the following incentive compatibility (IC) condition:[10]

$$\Pi_t + ra \geq (I - a) + w_t\ell + T_t. \tag{2.5}$$

Thus, what an agent earns from investing in a firm must be higher or equal than what she would obtain if she defaults with the loan and becomes a worker instead. This condition defines

[10]Taxes are collected on the basis of agents' wealth at the beginning of period t. Thus, if an agent decides to default and loses his wealth, she still has to pay taxes. That is the reason why τ_t does not appear in the IC condition. Additionally, the IC condition can be written in terms of debt $d = I - a$ as follows:

$$p_t R - rd \geq d + w_t\ell + T_t$$

a minimum wealth to obtain a loan, $\hat{a}_t \equiv \hat{a}(b_t, \Gamma_t)$. In the rest of the paper, I refer to $\hat{a}_t$ as the *minimum collateral*, which is implicitly defined by:

$$\hat{a}_t = I - \frac{p_t R - w_t \ell - T_t}{1+r}, \tag{2.6}$$

where p_t and w_t depend on the fraction of entrepreneurs, e_t, which is a function of the wealth distribution (Γ_t) and transfer rate (b_t). Additionally, total transfers T_t depend on the fraction of workers, $1 - e_t$, and the transfer rate. As a result, the minimum collateral becomes a function of (b_t, Γ_t).

2.6 Timing of Individual Decisions

Consider momentarily a discrete time model where the length of a period is $\Delta > 0$. Figure 2 illustrates the three sequential individual decisions that take place in a timeframe Δ.

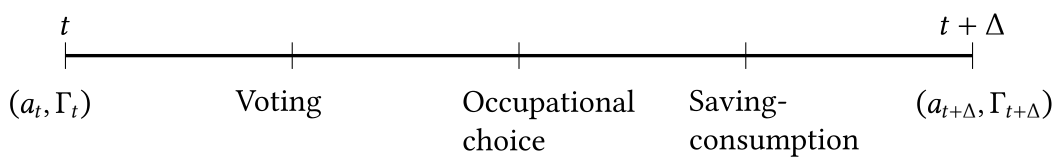


Figure 2: Timing of individual decisions

2.6.1 Voting

At the beginning of each period t, each agent votes for her preferred transfer rate after observing her wealth (a_t) and the wealth distribution (Γ_t). Each individual chooses her preferred transfer rate to maximize her current disposable income and before making her occupational choice.[11] Therefore, citizens define their political preferences while taking into account the effects on credit constraints and on their occupational decisions. The voting process results in an equilibrium transfer rate, b_t. The political process is described in detail in Section 5.

[11] A more standard alternative would be that agents want to maximize their utility from consumption. Lemma 1 shows that agents consume a fraction of their disposable income every period. Hence, it is equivalent to defining individual preferences in terms of the disposable income or in terms of the utility of consumption.

2.6.2 Occupational Choice

After observing the equilibrium transfer rate, agents choose between becoming workers or entrepreneurs. Investment decisions are limited by the minimum collateral, $\hat{a}_t$. The occupational decisions are characterized in Section 3.1.

2.6.3 Saving-Consumption

After agents have defined their occupations, they decide on consumption and savings to maximize their future discounted utility but without anticipating the future join evolution of policies and distributions. In particular, agents solve their Bellman equation thinking that the current transfer rate (b_t) will remain fixed in the future. Individual savings decisions define the future wealth distribution $\Gamma_{t+\Delta}$.

In sum, I make two important assumptions regarding voting and economic decisions. First, agents vote to maximize their current disposable income before making their occupational decisions. Second, when agents decide on consumption and savings they think that the current policy will remain unchanged in the future.

The alternative to these behavioral assumptions is to work with the fully-rational equilibrium in which agents predict the future evolution of all endogenous variables. So far, the literature that deals with this type of equilibrium has relied on numerical solutions (e.g. Krusell and Rios-Rull, 1999). To my knowledge, there are no theoretical results for this type of equilibrium in a model with heterogeneous agents that incorporates a political dimension. I overcome these technical challenges by relying on the two aforementioned behavioral assumptions. The model remains sufficiently tractable to allow for a sharp characterization of the transition dynamics (see Section 6).

3 Equilibrium

3.1 Occupational Choice

An agent decides to invest in a firm and to become an entrepreneur if the following occupational constraint (OC) is satisfied:

$$\Pi_t \geq w_t \ell + T_t, \tag{3.1}$$

Thus, an agent starts a firm only when the profits she receives are larger than her labor income plus social benefits. When this condition is binding, it gives rise to a wealth threshold that defines the first agent that is willing to start a firm: $\tilde{a}_t \equiv \tilde{a}(b_t, \Gamma_t)$. I refer to $\tilde{a}_t$ as the *occupational threshold.*

Occupational choice is determined by comparing the minimum collateral ($\hat{a}_t$) and the occupational threshold ($\tilde{a}_t$). Figure 3 illustrates occupational choice. In Case 1, the incentive compatibility constraint binds. Thus, the minimum collateral to get credit ($\hat{a}$) defines the first agent that starts a firm. Agents with assets $a \in (\tilde{a}, \hat{a})$ would like to become entrepreneurs. However, because they are excluded from the credit market, they cannot finance a firm and must become workers. In Case 2, the occupational constraint binds, so $\tilde{a}$ defines the first agent that becomes an entrepreneur. Individuals with assets $a \in (\hat{a}, \tilde{a})$ could obtain credit to finance a firm, but they prefer to become workers instead.

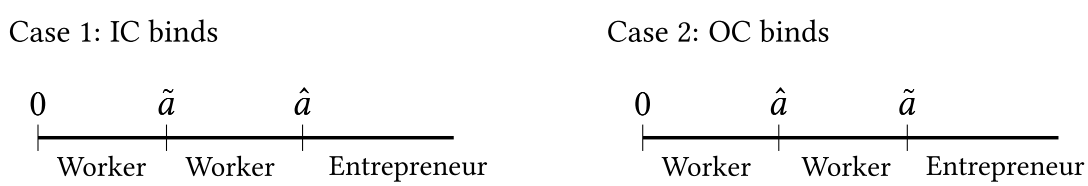


Figure 3: Occupational choice.

To sum up, the asset threshold that defines occupational choice is given by: $a_t^o \equiv \max\{\hat{a}_t, \tilde{a}_t\}$. I refer to a_t^o as the *effective occupational threshold.* In any given period, agents with $a < a_t^o$ become workers, while the rest become entrepreneurs.[12] Therefore, the endogenous fraction of

[12]Occupational choice works is in a similar way to that of Buera and Shin (2013). In their model, agents differ in

entrepreneurs is given by: $e_t = 1 - \Gamma_t(a_t^o)$.

Taking into account occupational choice, the individual problem is written compactly as follows:

$$
\begin{aligned}
&\max_{\{c_t\}_{t=0}^{+\infty}} \left\{ \int_0^\infty e^{-\rho t} \log(c_t) dt \right\} \\
&s.t. \qquad \dot{a}_t = y_t(a) - c_t, \\
&\qquad\qquad a_t^o = \max\{\hat{a}_t, \tilde{a}_t\}, \\
&\qquad\qquad a \geq \underline{a},
\end{aligned}
$$

where $y_t(a) = (r - \tau_t)a + (w_t\ell + T_t) \cdot \mathbb{1}_{a<a_t^o} + \Pi_t \cdot \mathbb{1}_{a\geq a_t^o}$ is the disposable income of an agent with assets a.

3.1.1 The Occupational Condition and the Incentive Compatibility Constraint

For the rest of the paper, it is useful to define the following occupational and incentive compatibility functions:

$$OC(e, b) = p(e)R - rI - w(e)\ell - bY(e), \tag{3.2}$$

$$IC(e, b) = p(e)R - rI - w(e)\ell - bY(e) - [I - (1 + r)\Gamma^{-1}(1 - e)], \tag{3.3}$$

where $\Gamma^{-1}(\cdot)$ denotes the inverse of the cumulative wealth distribution function. Thus, $\Gamma^{-1}(1 - e)$ gives the occupational threshold, a^o, that makes the fraction of entrepreneurs equal to e when the wealth distribution is Γ. In Section C.1 in the Appendix, I study the theoretical properties of

their wealth and entrepreneurial ability. Individuals with a given ability choose to become entrepreneurs only if they are wealthy enough to overcome an exogenous collateral constraint. Thus, there is a level of wealth above which agents with a certain ability prefer to start a firm. On the other hand, in my model all agents are equally productive and their occupational decision depends only on their wealth and the current state of the economy (Γ_t, b_t). As in Buera and Shin (2013), there is wealth threshold above which agents are willing to become entrepreneurs $(\tilde{a}_t)$, but there is also an endogenous minimum collateral that limits entrepreneurial decisions $(\hat{a}_t)$.

both the OC and IC functions in detail.

Figure 4 illustrates the properties of $OC(e, b)$ and $IC(e, b)$ for a fixed b and as a function of the fraction of entrepreneurs, e. Throughout the paper, I refer to Figure 4 as the *OC-IC diagram*. To simplify the analysis, I define the following functions:

$$h(e) \equiv p(e)R - w(e)\ell - bY(e) \tag{3.4}$$

$$x(e) \equiv (1 + r)[I - \Gamma^{-1}(1 - e)] \tag{3.5}$$

Thus, the occupational constraint function becomes: $OC(e) = h(e) - rI$, while the incentive compatibility is: $IC(e) = h(e) - x(e)$. The bold dotted line in Figure 4 represents the horizontal line at rI. The figure takes as given the transfer rate and the wealth distribution, therefore I omit the dependence on (b, Γ).

The difference between the solid line ($h(e)$) and the dotted line (rI) corresponds to $OC(e)$. When both lines intersect, the occupational constraint binds. The fraction of entrepreneurs at the intersection is denoted by $\tilde{e}$ and is the maximum fraction of entrepreneurs that is sustainable by the economy given a set of parameters and a wealth distribution Γ. The occupational threshold at that point is $\tilde{a} = \Gamma^{-1}(1 - \tilde{e})$.[13]

The distance between the solid line ($h(e)$) and the dashed line ($x(e)$) corresponds to $IC(e)$. The intersection of both lines happens at $\hat{e}$ which is the fraction of entrepreneurs that makes the IC constraint binding. The minimum collateral is then given by: $\hat{a} = \Gamma^{-1}(1 - \hat{e})$. In the figure, $\tilde{e} > \hat{e}$ and so $\hat{a} > \tilde{a}$. Thus, the occupational threshold is given by $a^o = \hat{a}$.

It is important to highlight that is not always the case that $\hat{a} > \tilde{a}$ and that the evolution of both thresholds depends on the transition dynamics of Γ. For instance, suppose that the wealth distribution is shifting right over time (in the First Order Stochastic Dominance (FOSD) sense). Then,

[13]The fraction of entrepreneurs $\tilde{e}$ that solves $OC(e) = 0$ is unique and does not depend on Γ (see Section C.1 in the Appendix). However, note that the wealth level above which agents are willing to become entrepreneurs, $\tilde{a}$, does depend on the distribution.

$x(e)$ also moves right over time. Because $h(e)$ does not change when Γ changes, the intersection of $h(e)$ and $x(e)$ shifts right. Eventually, the intersection happens at some $\hat{e} \geq \tilde{e}$, which implies that $\hat{a} \leq \tilde{a}$, and thus, the occupational condition becomes binding from that point onwards ($OC = 0$).

Finally, e' is the fraction of entrepreneurs at which $OC = IC$. The intersection of both functions always happens at $a^o = \frac{I}{1+r}$, which is the maximum minimum collateral sustainable in the economy. That is, $\hat{a}$ can be at most $\frac{I}{1+r}$, otherwise the OC always binds. Therefore, *the maximum sustainable transfer rate* given a wealth distribution Γ is:

$$\bar{b}(\Gamma) = \frac{p(e')R - rI - w(e')\ell}{Y(e')}. \tag{3.6}$$

where $e' = 1 - \Gamma(I/(1+r))$. Thus, social benefits as a fraction of GDP are restricted to be in the interval $[-\underline{b}, \bar{b}(\Gamma)]$.

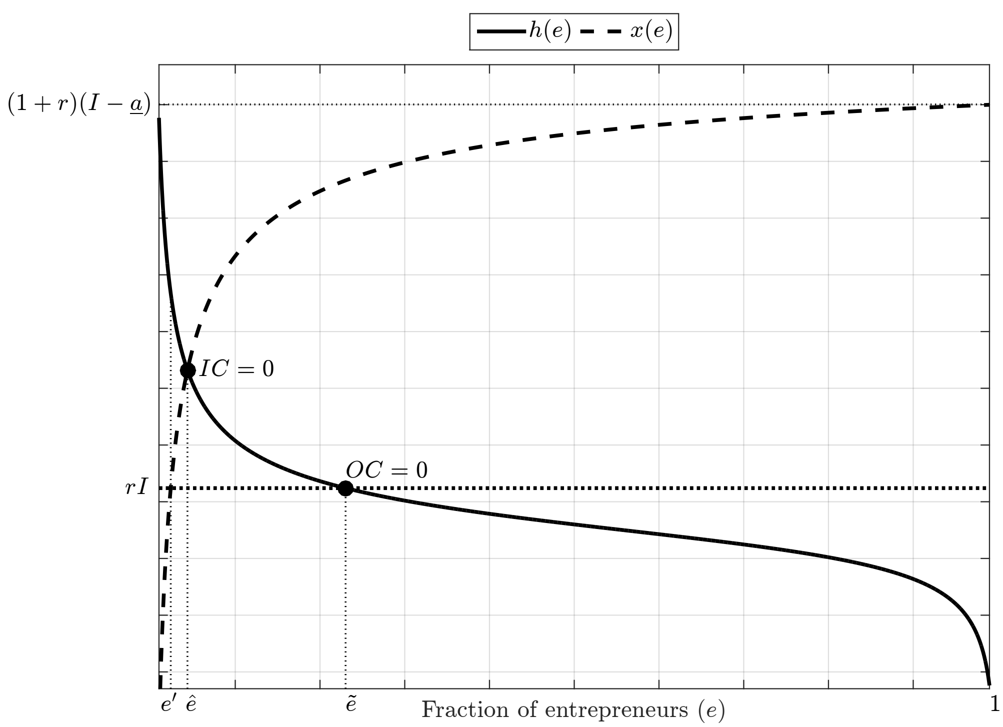


Figure 4: The OC-IC diagram.

3.2 Consumption and Saving Decisions

Lemma 1 makes use of two assumptions to derive close form solutions for the consumption policy function, $c_t(a)$ and the savings policy function, $s_t(a)$: i) agents have a logarithmic utility, and ii) agents do not predict the future evolution of transfer rates and distributions. Every period, agents consume and save a fraction of their disposable income, $y_t(a)$. The saving rate, θ_t, depends on the fixed interest rate r and the discount factor ρ, but also on the tax rate τ_t which is a function of the transfer rate b_t. Thus, the evolution of saving decisions depends on the endogenous evolution of the size of the welfare state.

The saving rate is a positive function of the ratio: $\frac{r-\tau_t}{\rho}$. Intuitively, when the "effective rate" at which agents can save (i.e. $r-\tau_t$) is larger than the discount factor, they want to save a positive fraction of their disposable income. Otherwise, agents dissaccumulate assets.

Lemma 1 *The optimal consumption and savings policy functions are linear functions of disposable income $y_t(a)$:*

$$c_t(a) = (1-\theta_t)\cdot y_t(a), \tag{3.7}$$

$$s_t(a) = \theta_t \cdot y_t(a), \tag{3.8}$$

where the saving rate is: $\theta_t = \left(1 - \frac{\rho}{r-\tau_t}\right)$.

Integrating the saving policy across all agents gives an expression for the evolution of aggregate wealth:

$$\dot{A}_t = \theta_t[rA_t + w_t\ell\cdot(1-e_t) + \Pi_t\cdot e_t]. \tag{3.9}$$

Thus, the change in aggregate wealth is a fraction θ_t of total income in the economy which is the sum of: interest income rA_t, labor income $w_t\ell\cdot(1-e_t)$, and firms' profits $\Pi_t\cdot e_t$.

3.3 Equilibrium Definition

I restrict attention to equilibria that satisfy the Markov property. Specifically, the equilibrium transfer rate (b_t) in each period is only a function of the current wealth distribution (Γ_t). In Section 5, I describe the political process that maps the wealth distribution into the equilibrium transfer rate. In particular, I show that under the behavioral assumptions described in Section 2.6, b_t becomes a function of the cumulative wealth distribution, Γ_t. I refer to that function as the *Political Equilibrium* (PE) condition: $b_t = P(\Gamma_t)$.

The dynamics of Γ_t are described by the *Kolmogorov Forward* (KF) equation : $d_t\Gamma_t(a) = H(\Gamma_t, b_t)$.[14] For any wealth distribution and transfer rate, the function $H(\cdot)$ gives the future change of the cumulative distribution at any level of assets. This function is generated by individuals' consumption-saving decisions presented in Lemma 1.

The PE and KF equations describe the joint evolution of the wealth distribution and transfer rate over time. The PE condition defines a "political mapping" from Γ_t to the equilibrium size of the welfare state, b_t. On the other hand, the KF equation maps the current size of the welfare state and distribution to the future distribution based on economic decisions. The main feature of the model is the endogenous feedback between the wealth distribution and the welfare state over time. Figure 5 illustrates this dynamic relationship.

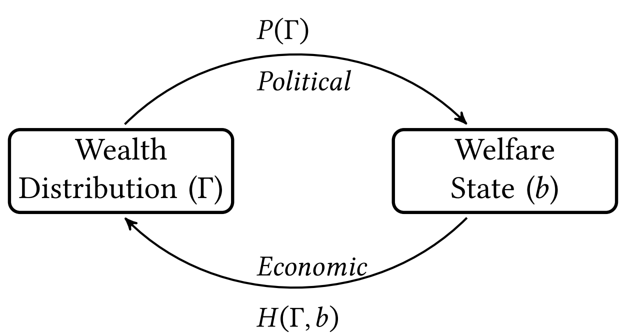


Figure 5: Joint evolution of Γ and b.

Given some initial wealth distribution Γ_0, the evolution of the economy is characterized by

[14]In the rest of the paper, the partial derivative in terms of some variable x, $\frac{d}{dx}(\cdot)$ is denoted by $d_x(\cdot)$.

the following set of equations:

$$\rho v_t(a) = \max_c \left\{ \log(c) + d_a v_t(a)\left(y_t(a, b_t) - c\right) + d_t v(a) \right\} \tag{3.10}$$

$$d_t\Gamma_t = H(\Gamma_t, b_t) \tag{3.11}$$

$$a_t^o = \max\left\{ \hat{a}(\Gamma_t, b_t), \tilde{a}(\Gamma_t, b_t) \right\} \tag{3.12}$$

$$b_t = P(\Gamma_t) \tag{3.13}$$

$$\tau_t = \frac{b_t Y_t \left(1 - \Gamma_t(a_t^o)\right)}{A_t} \tag{3.14}$$

Therefore, a *political equilibrium* is such that: i) agents solve the Hamilton Jacobi Bellman (HJB) equation (3.10) by taking as given the price of capital p_t and the wage rate w_t, ii) the evolution of the wealth distribution is given by the KF equation (3.11), iii) occupational choice is determined by the effective occupational threshold, a_t^o, as defined by equation (3.12), iv) at any given period, the transfer rate b_t is defined by the PE condition (3.13), v) the tax rate keeps a balanced budget as defined by equation (3.14), and vi) prices are given by: $p_t = \alpha Z(Re_t)^{\alpha-1}(\ell(1-e_t))^{1-\alpha}$ and $w_t = (1-\alpha)Z(Re_t)^{\alpha}(\ell(1-e_t))^{-\alpha}$, where the fraction of entrepreneurs is $e_t = 1 - \Gamma_t(a_t^o)$.

3.4 The Evolution of the Wealth Distribution

The following lemma provides an explicit expression for the KF equation (3.11).[15]

Lemma 2 *The evolution of the cumulative wealth distribution $\Gamma_t(a)$ is characterized by the Kolmogorov Forward (KF) equation:*

$$d_t\Gamma_t(a) = -\Gamma_t(a_t^o) \cdot s_t(a) d_a\Gamma_t(a) - \left(1 - \Gamma_t(a_t^o)\right) \cdot s_t(a) d_a\Gamma_t(a). \tag{3.15}$$

[15] Note that $-\Gamma_t(a_t^o) \cdot s_t(a) d_a\Gamma_t(a)$ does not cancel out in equation (3.15). Recall that $s_t(a) = \theta_t y_t(a)$, where the disposable income is $y_t(a) = (r - \tau_t)a + (w_t\ell + T_t)\mathbb{1}_{a<a_t^o} + \Pi_t \mathbb{1}_{a \geq a_t^o}$. Thus, the KF equation can be written as:

$$d_t\Gamma_t(a) = \begin{cases} -\Gamma_t(a_t^o) \cdot \theta_t[(r - \tau_t)a + w_t\ell + T_t] \cdot d_a\Gamma_t(a) & \text{if } a < a_t^o \\ -\left(1 - \Gamma_t(a_t^o)\right) \cdot \theta_t[(r - \tau_t)a + \Pi_t] \cdot d_a\Gamma_t(a) & \text{if } a \geq a_t^o \end{cases}$$

The first term of the KF equation captures the inflow and outflows due to continuous movements in the wealth of workers, while the second term does the same for entrepreneurs. Equation (3.15) implicitly defines the function $H(\cdot)$ in equation (3.11) that maps the current policy b_t and distribution Γ_t to the future wealth distribution.

Section C.3 in the Appendix, illustrates a one-period shift of the wealth distribution when agents save a positive fraction of their income, $s_t(a) > 0$. In particular, the wealth distribution always shifts in the FOSD sense over time.[16]

3.5 Stationary Equilibrium

In this section, I restrict attention to non-degenerate stationary wealth distributions.[17] A stationary equilibrium is a political equilibrium satisfying:

$$d_t\Gamma_t = 0,\ A_t = A^*,\ w_t = w^*,\ p_t = p^*,\ b_t = b^*,\ \tau_t = \tau^*,\ a_t^o = a^{o*} \quad \text{for all } t. \tag{3.16}$$

Imposing these restrictions in equations (3.10) to (3.14) yield the following Lemma.[18]

Lemma 3

1. *There is a unique stationary tax rate,* $\tau^* = r - \rho$.

2. *The stationary wealth distribution,* Γ^**, is non-unique. There is a set of stationary distributions*

[16]This property comes from two facts: i) there is no uncertainty, and ii) all agents save a fraction of their disposable over time, where the saving rate is the same across agents in a given period.

[17]In particular, when I study transitions dynamics in Sections 4 and 6, I restrict the initial distribution to the set of continuously differentiable distributions with support in $[\underline{a}, +\infty)$. In general, this set of initial distributions generates stationary wealth distributions that are non-degenerate. There are some few exceptions in which the wealth distribution collapses to $\underline{a}$, in which case the stationary tax rate can be different to the one identified in Lemma 3 (see Section 6 for a complete discussion).

[18]Note that $a_t^o = \max\left\{\hat{a}(\Gamma_t, b_t), \tilde{a}(\Gamma_t, b_t)\right\}$ and $b_t = P(\Gamma_t)$. Thus, $a_t^o = \max\left\{\hat{a}(\Gamma_t, P(\Gamma_t)), \tilde{a}(\Gamma_t, P(\Gamma_t))\right\}$ so a_t^o can be written as a function of Γ_t. In Lemma 3, I denote this function by $\Lambda(\cdot)$.

that solves the system:

$$r - \rho = \frac{b^* \Gamma^* (a^{o*}) Y(\Gamma^*)}{A^*}, \tag{3.17}$$

$$a^{o*} = \Lambda(\Gamma^*), \tag{3.18}$$

$$b^* = P(\Gamma^*). \tag{3.19}$$

The first item of Lemma 3 states that there is a unique stationary tax rate τ^*. Recall that agents save a fraction θ_t of their disposable income every period. If the aggregate productivity Z is not changing, then the only way in which the wealth distribution can remain constant is if agents do not save, i.e. if $\theta_t = 0$.[19] This condition implies that the equilibrium tax rate must be equal to $r - \rho$ in the long-run.

The second item states that, given a set of parameters other than the initial wealth distribution, there is a set of stationary wealth distributions. However, given some initial wealth distribution, the economy may reach at most one stationary distribution in the long-run. Krusell and Rios-Rull (1996, 1999) obtained a similar result regarding the non-uniqueness of the stationary wealth distribution after introducing politics to the standard neoclassical model.

Two key questions arise from Lemma 3: i) whether the economy will converge to τ^* starting from any arbitrary wealth distribution, and ii) whether it will attain a stationary wealth distribution at all.

In Section 6, I characterize the transition dynamics and address these questions by studying the convergence properties of the model. In general, the political equilibrium converges to τ^* in the long-run. However, in some cases, the wealth distribution may diverge, in which case the tax rate goes to zero in the long-run. Even in these more complicated cases, it is possible to characterize the evolution of the size of the welfare state which always attains some stationary level. An important feature of the model is that the initial wealth distribution matters for the transition dynamics of the size of the welfare state and its steady state level.

[19] In the quantitative analysis in Section 7, the aggregate productivity Z can potentially follow an exogenous path over time. Thus, a necessary condition to have a stationary equilibrium is that Z remains constant in the long-run.

Finally, the steady state level of the rest of the variables such as prices depend on the stationary wealth distribution which is non-unique. Hence, there is a set of steady-state prices, transfer rates, and occupational thresholds that satisfy (3.16).

4 Transition Dynamics: Exogenous Policy

This section characterizes the transition dynamics when the size the of the welfare state is fixed over time: $b_t = b$. Thus, the equilibrium of the economy is obtained by substituting $P(\Gamma_t) = b$ into equation (3.13). This serves as a useful starting point before analyzing the political equilibrium, where the evolution of b_t is chosen through repeated voting over time.

The timing of events in period t is as follows. 1) Agents observe their assets a and the wealth distribution at the beginning of the period (Γ_t). 2) Banks observe the wealth distribution and define the minimum collateral required for a loan, $\hat{a}_t$. 3) Based on $\hat{a}_t$ and the occupational threshold $\tilde{a}_t$, agents make their occupational choice according to: $a_t^o = \max\left\{\hat{a}_t, \tilde{a}_t\right\}$. 4) Workers and entrepreneurs make their saving-consumption decisions.

The following proposition characterizes the transition dynamics given an initial wealth distribution Γ_0 which is continuously differentiable and has support in $[\underline{a}, +\infty)$.

Proposition 1 *Consider an initial wealth distribution Γ_0. Define the initial tax rate $\tau_0 \equiv \tau(\Gamma_0)$ and the initial fraction of entrepreneurs $e_0 \equiv 1 - \Gamma_0\left(a^o(\Gamma_0)\right)$. Further, restrict Γ_0 to the set of distributions that satisfy $\tau_0 \leq r - \rho$ and $e_0 \geq \frac{\alpha}{2}$. Consider the following three cases for which the transitions dynamics can be completely characterized:*

1. *If $r - \rho < 0$ and $b < 0$, then:*
    - $d_t\tau_t \geq 0$ *and* $\tau^* = r - \rho$.
2. *If $r - \rho > 0$ and $b < 0$, then:*
    - $d_t\tau \geq 0$ *and* $\tau^* = 0$.

3. *If $r-\rho>0$ and $b>0$, then:*

   - *$d_t\tau \leq 0$ and $\tau^* = 0$.*

All three cases satisfy: i) $d_t\Gamma_t(a) \leq 0\ \forall a$, ii) $d_t a_t^o \geq 0$, iii) $d_t k_t \geq 0$, iv) $\exists \tilde{t} > 0 : k_t = R\tilde{e},\ \forall t \geq \tilde{t}$, and v) $d_t A_t \geq 0$. The steady state fraction of entrepreneurs, $\tilde{e}$, solves $OC(\tilde{e}, b) = 0$ in equation (3.2).

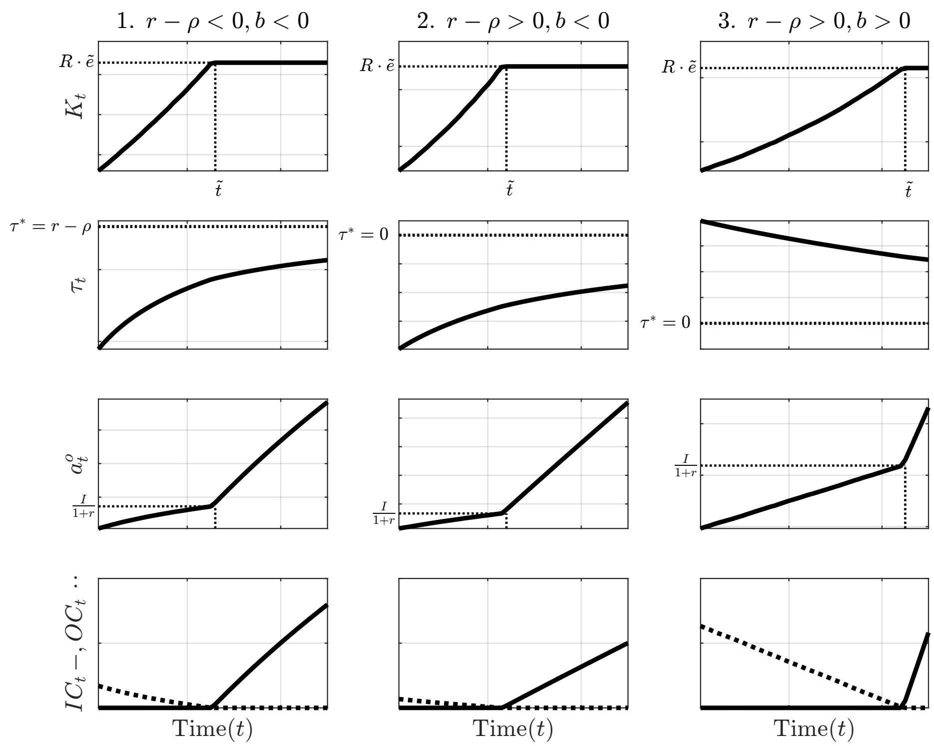


Figure 6: Transition dynamics, fixed transfer rate ($b_t = b$).

Figure 6 illustrates the three cases described in Proposition 1. The figure presents the transitions dynamics of capital (K_t), tax rate (τ_t), occupational threshold (a_t^o), the incentive compatibility (IC_t, solid line) and occupational constraint (OC_t, dotted line).

First, in all cases, capital increases over time and attains a steady state level after a finite number of periods $\tilde{t} > 0$.

Second, the tax rate converges to $r-\rho$ only in the first case. In the remaining two cases, τ converges to zero in the long-run which implies that agents keep saving indefinitely ($\theta^* = \frac{r-\rho}{r} >$

0), and thus, the wealth distribution diverges.

Finally, the incentive compatibility constraint initially binds, meaning that occupational choice is restricted by credit conditions ($\hat{a} > \tilde{a}$). Eventually, when the economy accumulates enough wealth, the occupational constraint binds and credit constraints do not limit occupational choice anymore ($\hat{a} < \tilde{a}$).

In order to understand Proposition 1, consider momentarily a discrete time model where the length of a period is $\Delta > 0$. In what follows, I explain the results stated in items i) to v).

First, suppose that $\tau_0 < r - \rho$, then i) states that: $\Gamma_t(a) \geq \Gamma_{t+\Delta}(a), \forall a$. Thus $\Gamma_{t+\Delta}(a)$ FOSD $\Gamma_t(a)$. That is, the cumulative wealth distribution shifts to the right. The intuition is that whenever $\tau_t < r - \rho$, the saving rate out of disposable income is positive, $\theta_t \geq 0$. Therefore, $s_t(a) \geq 0, \forall a, \forall t$. Since agents are saving a positive fraction of assets each period, the wealth of each agent is weakly increasing over time and the cumulative wealth distribution shifts right. This implies that aggregate wealth is increasing over time as stated by item v).

Second, to understand item ii), recall the minimum collateral condition that arises from the IC constraint:

$$\hat{a}_t = \left(I - \frac{p_t R - w_t \ell - bY_t}{1+r} \right). \tag{4.1}$$

Initially, when the economy has low wealth the IC constraint binds. Thus, the evolution of occupational choice is determined by the minimum collateral, i.e. $a_t^o = \hat{a}_t$. Because $\Gamma_{t+\Delta}$ FOSD Γ_t, the mass of entrepreneurs for a given $\hat{a}$ increases over time, i.e. physical capital goes up. Thus, the price of capital decreases, while the wage rate and output increase.[20] To satisfy (4.1), banks tighten credit conditions over time, i.e. $\hat{a}_{t+\Delta} > \hat{a}_t$. When the economy has accumulated enough wealth (at $t = \tilde{t}$), the OC constraint binds so the fraction of entrepreneurs equals $\tilde{e}$ from that point onwards (item iv)).[21]

[20]The effect of increasing the fraction of entrepreneurs on output is ambiguous, because capital increases but aggregate labor goes down. In the proof of item ii) of Proposition 1, I show that Assumption 1 on $\underline{b}$ is a sufficient condition to have that when e increases output also increases.

[21]Note that at $t = \tilde{t}$, when the OC and IC functions intersect, the economy reaches the maximum sustainable minimum collateral, $\frac{I}{1+r}$. Because the OC becomes binding from that point onwards, there is a kink in the occupational threshold curve (a_t^o). Formally, this can be seen in equation (A.22) in the Appendix.

Finally, the evolution of capital K_t over time depends on the evolution of the fraction of entrepreneurs: $e_t = 1 - \Gamma_t(a_t^o)$. From the previous discussion, there are two opposite effects over time: I) $\Gamma_t(a)$ shifts right in the FOSD sense, and thus, e_t increases for a given a_t^o, and II) a_t^o increases which reduces e_t for a given distribution $\Gamma_t(a)$. Proposition 1 shows that the distributional effect I) dominates, and thus, $e_{t+\Delta} \geq e_t$. As a result, capital increases over time $K_{t+\Delta} \geq K_t$.

The dynamics presented in Figure 6 can be understood through the lens of the OC-IC diagram presented in Section 3.1.1. Section C.2 in the Appendix provides a complete discussion.

5 Political Process

This section is organized as follows. In Section 5.1, I start by studying the basic comparative statics regarding the transfer rate. In Section 5.2, I characterize the individual preferences for the transfer rate—the size of the welfare state—that are induced via the economic equilibrium. In Section 5.3, I present the condition that defines the equilibrium size of the welfare state after aggregating the individual preferences through the political process. For space considerations, I present a detailed description of the political process in Section B.1 in the Appendix.

5.1 The Transfer Rate: Comparative Statics

Before I study the individual preferences for the welfare state, the following lemma describes the basic comparative statics regarding the transfer rate.

Lemma 4 *Given some wealth distribution Γ_t, a marginal increase in the transfer rate, b, leads to: i) an increase of the minimum collateral $\hat{a}$, ii) an increase of the occupational threshold $\tilde{a}$, iii) a decrease of the wage rate w, and iv) and increase of the price of capital p.*

When the transfer rate increases, the first order effect is that transfers to workers increase making the IC and OC constraints more binding. Thus, the effective occupational threshold (a^o) increases reducing the fraction of entrepreneurs, and thus, the production of physical capital. As a result, there are two important second order effects: the price of capital increases, while the wage rate decreases. Both effects decrease to some extent the effective occupational threshold.[22] However, the first order effect dominates and so, both the minimum collateral ($\hat{a}$) and the occupational threshold ($\tilde{a}$) go up when b increases.

Overall, when social benefits to workers increase, the opportunity cost of starting a firm goes up, decreasing entrepreneurship. This result captures the recent empirical findings of a large body of literature that provides evidence of a negative relationship between social benefits and entrepreneurship (Audretsch et al., 2022; Solomon et al., 2022, 2021; Song et al., 2020). Also, earlier studies include Henrekson (2005); Hessels et al. (2006, 2008); Koellinger and Minniti (2009).

5.2 Individual Preferences

In this section, I study the individual preferences for the size of the welfare state as measured by the transfer rate, b. At the beginning of each period, agents observe their assets a and the wealth distribution, Γ. Then, each agent defines her preferred transfer rate, denoted by $b(a, \Gamma)$, to maximize her current disposable income.[23] Individuals choose $b(a, \Gamma)$ before making consumption-saving decisions but anticipating the effects on occupational choice. Formally,

$$b(a;\Gamma) = \underset{b\in[-\underline{b},\ \overline{b}(\Gamma)]}{\arg\max} \left\{ (r-\tau(b,\Gamma))a + (w(b,\Gamma)\ell + T(b,\Gamma)) \cdot \mathbb{1}_{a<a^o(b,\Gamma)} + \Pi(b,\Gamma) \cdot \mathbb{1}_{a\geq a^o(b,\Gamma)} \right\}. \quad (5.1)$$

The minimum possible transfer rate $-\underline{b}$ satisfies Assumption 1, while $\overline{b}(\Gamma)$ corresponds to the maximum sustainable transfer rate as defined by equation (3.6).

Lemma 5 *The preferred transfer rate function $b(a, \Gamma)$ is as follows:*

[22] Also, when the fraction of entrepreneurs decreases, total output (Y) goes down. Therefore, transfers to workers $T = b \cdot Y$ go down for a given b, decreasing the effective occupational threshold to some extent.

[23] A more standard alternative would be that agents want to maximize their utility from consumption. As shown in Lemma 1, agents consume a fraction of their disposable income every period. Hence, it is equivalent to defining individual preferences in terms of the disposable income or in terms of the utility of consumption. I opt for the first alternative because it simplifies the aggregation of preferences keeping the model tractable.

$$b(a;\Gamma) = \begin{cases} \psi^1(a;\Gamma) & \text{if } a < \hat{a}(\text{-}\underline{b},\Gamma), \\ \psi^2(a;\Gamma) & \text{if } a \in [\hat{a}(\text{-}\underline{b},\Gamma), \overline{a}(\Gamma)], \\ \psi^3(a;\Gamma) & \text{if } a \geq \overline{a}(\Gamma), \end{cases} \tag{5.2}$$

where the functions $\psi^j(a;\Gamma)$, $j \in \{1,2,3\}$, *are continuous in assets and satisfy:*

$d_a\psi^1 \leq 0$, $d_a\psi^2 > 0$, $d_a\psi^3 \leq 0$, $\psi^2(\hat{a}(\text{-}\underline{b};\Gamma)) < \lim_{a\to\hat{a}(\text{-}\underline{b},\Gamma)^-} \psi^1(a;\Gamma)$ *and* $\lim_{a\to+\infty} \psi^3(a;\Gamma) = \text{-}\underline{b}$. *Further,* $\overline{a}(\Gamma) \in (\hat{a}(\text{-}\underline{b},\Gamma), \tilde{a}(\overline{b},\Gamma)]$.

Figure 7 depicts the preferred transfer rate function, $b(a;\Gamma)$ (black solid line). I classify agents into three classes according to their preferences and occupational prospects: the Working class, the Emerging class, and the Incumbent class. Overall, there is an endogenous conflict regarding the size of the welfare state between the three classes.

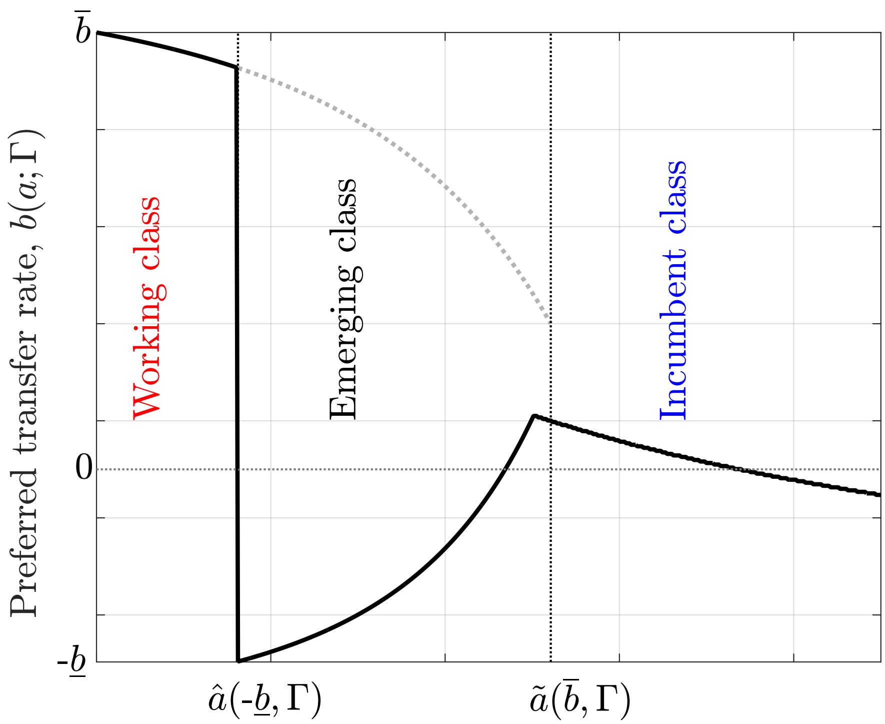


Figure 7: The preferred transfer rate function, $b(a;\Gamma)$.

The three classes are defined by two endogenous thresholds that depend on the current wealth distribution: i) the lowest possible minimum collateral to get credit, $\hat{a}(-\underline{b}, \Gamma)$, and ii) the maximum occupational threshold, $\tilde{a}(\overline{b}, \Gamma)$. The first threshold is attained when policies are the most favorable for entrepreneurship, i.e. $b = -\underline{b}$. The second threshold is defined by the maximum sustainable transfer rate $\overline{b}$, i.e. when the welfare state reaches it largest possible size. In the interval $[-\underline{b}, \overline{b})$, the IC constraint binds while the OC condition does not. Thus, $\Pi > w\ell + T$, and so it is profitable to invest in a firm and become an entrepreneur. In what follows, I describe the preferences of each class.

Firstly, individuals with less assets than $\hat{a}(-\underline{b}, \Gamma)$ must become workers regardless of the size of the welfare state. Thus, agents from the working class in general advocate for high social benefits. They trade-off higher transfers at the cost of lower wages and higher taxes. Wealthier agents from this class must finance a larger fraction of social benefits through taxes. Thus, $b(a; \Gamma)$ is decreasing in assets within the working class.

Secondly, individuals with assets $a \in [\hat{a}(-\underline{b}, \Gamma), \overline{a}(\Gamma))$ may be able to start a firm or not depending on the size of the welfare state (the Emerging class). Agents from this class trade-off three effects. A higher transfer rate increases the price at which they can sell their physical capital (*price effect*), but reduces their capital income (*capital income effect*), and increases the minimum collateral (*collateral effect*).

The poorest agents from the emerging class in general prefer business-supporting policies (negative b) , which relax credit constraints allowing them to start their own businesses. Thus, the collateral effect dominates for these agents.[24] Wealthier agents from this class prefer less pro-business policies to prevent the poorer agents from entering the market, maintaining low competition and high prices. Therefore, the price effect explains why $b(a; \Gamma)$ is in general increasing in assets for the emerging class. Overall, the agents from the emerging class may be willing to support a pro-business policy (negative b) sacrificing social benefits, but aspiring to become entrepreneurs at the end of the period. I refer to this behavior as *aspirational voting* which will play a key role in defining the equilibrium size of the welfare state (see next section).

[24]In particular, these agents prefer a transfer rate that allows them to exactly satisfy the minimum collateral requirement, i.e. an agent with wealth a chooses a b such that $a = \hat{a}(b, \Gamma)$

The disposable income as function of b for agents from the emerging class is not single peaked. In particular, there are two peaks (see Figure 14 in the Appendix).[25] When the transfer rate is such that $a < a^o$, they become workers, and thus, they demand high social benefits. However, when $a \geq a^o$, they support more pro-business policies. The gray dotted line in Figure 7 corresponds to the preferred transfer rate of emerging agents if they have to become workers. The black solid line represents their most preferred policy, i.e. such that they can start a firm.

Finally, agents that have more assets than $\tilde{a}(\bar{b}, \Gamma)$ can start a firm regardless of the size of the welfare state (Incumbent class). Agents from this class in general prefer less pro-business policies than the emerging class to prevent them from entering the capital market. Less wealthy agents from this class may be even willing to pay for social benefits to discourage agents to start a firm, and in this way, protect their businesses. Very wealthy agents from the incumbent class mainly care about their capital income, and thus, prefer lower social benefits. This explains why $b(a, \Gamma)$ is decreasing in assets within the incumbent class.

5.2.1 The Emerging and Incumbent Class: Related Literature

The finding of an emerging class that votes aspirationally for pro-business policies resembles the prospect of upward mobility hypothesis (POUM), which is one of the possible explanations proposed in the literature to explain why the poor may not support high redistribution (Benabou and Ok, 2001).[26] In my model, aspirational voting differs from the standard POUM hypothesis in that is not the poor but the middle class that is willing to support low redistribution, and not only because they want to become richer but also because they aspire to change occupation and join the business class.

The prediction of an incumbent class that is against "highly" pro-business policies shares similarities with the interest group theory proposed by Rajan and Zingales (2003) to explain why after World War I many countries remained financially undeveloped until 1980. According to

[25]In Section C.4 in the Appendix, I Illustrate how the preferred transfer rate function can be understood through the lens of the disposable income function.

[26]Related work also includes Acemoglu et al. (2018); Leventoğlu (2014); Benabou and Tirole (2006); Alesina and La Ferrara (2005); Leventoğlu (2005); Piketty (1995)

their theory, incumbents oppose financial development because it breeds competition (see also La Porta et al., 2000).[27]

5.3 The Equilibrium Size of the Welfare State

In each period, the equilibrium transfer rate maximizes a weighted measure of workers' and entrepreneurs' income:

$$\max_{b\in[-\underline{b},\bar{b}(\Gamma_t)]} \left\{ \mathcal{W}(b,\Gamma_t) \equiv w_t \ell(1-e_t) + \phi \Pi_t e_t \right\} \tag{5.3}$$

$$\text{s.t.} \quad e_t = 1 - \Gamma_t(a^o(b,\Gamma_t))$$

$$a_t^o = \max\{\hat{a}(b,\Gamma_t), \tilde{a}(b,\Gamma_t)\}$$

where $\phi > 1$ can be interpreted as a "political weight" which captures the political orientation of a "representative government", either more pro-worker or pro-business. In Section B.1 in the Appendix, I provide a political economy microfoundation for problem (5.3). I show that the problem can be rationalized by a probabilistic voting model à la Persson and Tabellini (2000). In that model, two candidate governments simultaneously announce their proposed size of the welfare state to maximize their share of votes.[28]

It is important to highlight that the political mechanism used to aggregate the individual preferences in not key for the qualitative properties of the evolution of the welfare state. The important idea captured by the political mechanism is that a representative government chooses spending on social benefits—without commitment and sequentially over time—while considering the citizens' interests. Thus, a "vote" means more broadly the *de facto* influence of an individual on the political aggregation process, but not literally her vote on elections. This interpretation

[27] Shleifer and Wolfenzon (2002) formalizes this hypothesis in a model of an entrepreneur going public under poor legal protection of outside shareholders. Rajan and Ramcharan (2011) find evidence that elites may restrict financial development to constrain access to finance of tenants and farmers.

[28] From Section 5.2, recall that preferences are not single-peaked, thus the median voter approach can give rise to cycling problems. To address this issue, the probabilistic voting creates a smooth mapping from policies to expected votes by introducing uncertainty about the outcome of elections.

also captures the fact that government spending decisions occur at a much higher frequency than elections. Additionally, in some countries certain groups may have a higher influence on policy decisions, which is captured in my model by the political weight, ϕ.

The idea of interests' aggregation should apply in general to different political systems. In fact, quoting Alesina and Rodrik (1994):

> "Even a dictator cannot completely ignore social demands, for fear of being overthrown. Thus, even in a dictatorship, distributional issues affecting the majority of the population will influence policy decisions".

In Lemma 6, I show that problem (5.3) has a unique solution, denoted as b_t. In particular, the objective function is strictly concave in e_t, so there is an "optimal" fraction of entrepreneurs, e^*, that solves the problem given a political weight, ϕ.

Lemma 6 *There is a unique transfer rate, b_t, that solves the government's problem* (5.3)*:*

$$1 - \Gamma_t(a^o(b_t, \Gamma_t)) = e^*, \tag{5.4}$$

where $e^* < \alpha$ is a function of aggregate productivity and the fixed parameters of the model.

Equation (5.4) corresponds to the *political equilibrium* (PE) condition that provides an implicit mapping from the wealth distribution to the equilibrium level of social benefits ($b_t = P(\Gamma_t)$). Equation (A.29) in the Appendix provides a more explicit expression for condition (5.4).[29] Lemma 6 gives the following corollary:

Corollary 1 *Define $\varphi = \frac{(\phi-1)(1-\alpha)}{1-\alpha(1-\phi)}$ and the political equilibrium output: $Y_{PE} = Z(Re^*)^\alpha(\ell(1-e^*))^{1-\alpha}$.*

1. *The effective occupational threshold is given by: $a^o_t = \Gamma_t^{-1}(1 - e^*)$.*

[29]In the theoretical analysis in this section the aggregate productivity Z is fixed over time, but in the quantitative analysis in Section 7 it can potentially follow an exogenous path. In that case, the optimal fraction of entrepreneurs is time-dependent, denoted as e^*_t.

2. *The occupational and incentive compatibility functions read as*

$$OC(b_t) = \varphi rI - b_t Y_{PE}, \tag{5.5}$$

$$IC(b_t) = \varphi rI - b_t Y_{PE} - [I - (1+r)\Gamma_t^{-1}(1-e^*)]. \tag{5.6}$$

3. *The equilibrium transfer rate b_t is given by:*

$$b_t = \begin{cases} \frac{(\varphi r - 1)I + (1+r)a_t^o}{Y_{PE}} & if\, a_t^o \leq \frac{I}{1+r}, \\ \overline{b} \equiv \frac{\varphi rI}{Y_{PE}} & if\, a_t^o > \frac{I}{1+r}. \end{cases} \tag{5.7}$$

Under the PE condition, the fraction of entrepreneurs is fixed to e^*. Thus, the OC function represented by equation (3.2) in Section 3.1.1 becomes a function only of b_t. When the OC binds ($OC(b_t) = 0$), the economy reaches the maximum sustainable transfer rate: $\overline{b} = \frac{\varphi rI}{Y_{PE}}$. The OC and IC intersect at $a_t^o = \frac{I}{1+r}$. Hence, when $a_t^o \leq \frac{I}{1+r}$, the transfer rate is determined by the IC constraint (5.6). Otherwise, the OC binds and b_t is equal to $\overline{b}$. In Section C.5 in the Appendix, I use a simple two dimensional graph to illustrate equation (5.7) and show how the transfer rate responds to exogenous changes in the wealth distribution.

5.3.1 The Welfare State and Aspirational Voting

The PE condition (5.4) captures the role of aspirational voters in a neat way. To illustrate this, suppose that at certain point in time the saving rate is positive, $\theta_t > 0$. Therefore, the wealth distribution shifts right in the FOSD sense. Agents become wealthier, and thus, the fraction of people that can start a firm increases. The PE condition becomes: $1 - \Gamma_t(a^o(b_t, \Gamma_t)) > e^*$.

To satisfy the PE condition, the government must increase the transfer rate, b_t, to increase the effective occupational threshold, a^o, and decrease the fraction of entrepreneurs. A more intuitive interpretation is that because more agents can start a firm, firms profits go down, reducing the attractiveness of entrepreneurship. As a result, aspirational voting becomes weaker and so agents demand higher social benefits.[30] Overall, aspirational voting plays a crucial role in determining the evolution of the welfare state.

6 The Evolution of the Welfare State

In this section, I provide a theoretical characterization for the evolution of the size of the welfare sate, as measured by the transfer rate, b_t. The political equilibrium is formally described by equations (3.10) to (3.14). The function that maps Γ_t to an equilibrium transfer rate b_t (denoted by $P(\cdot)$ in equation (3.13)) is implicitly given by equation (5.4). The dynamics can be sorted into four general cases described in Proposition 2: 1. (a), 1. (b), 2. (a), and 2. (b) Each case depends mainly on $r - \rho$ and the initial wealth distribution, Γ_0.

Proposition 2 *Consider an initial wealth distribution, Γ_0. Define the initial transfer rate $b_0 = P(\Gamma_0)$ and the initial tax rate $\tau_0 = \tau(\Gamma_0)$. Further, denote the long-run value of some variable x by $x^* = \lim_{t \to +\infty} x_t$. Then, the dynamics of the political equilibrium are as follows:*

1. *If $\tau_0 < r - \rho$, then: $d_t b_t \geq 0$, $d_t A_t \geq 0$, and $d_t a_t^o \geq 0$.*
   (a) *If $r - \rho < 0$, then: $b^* \in (\text{-}\underline{b}, 0)$, $\tau^* = r - \rho$, and $A^* \in (\underline{a}, +\infty)$.*
   (b) *If $r - \rho > 0$, there are two cases:*
      i. *$b^* \in (0, \bar{b}]$, $\tau^* = r - \rho$, and $A^* \in (\underline{a}, +\infty)$.*
      ii. *$b^* = \text{-}\bar{b}$, $\tau^* = 0$, and $A^* \to +\infty$.*
2. *If $\tau_0 > r - \rho$, then: $d_t b_t \leq 0$, $d_t A_t \leq 0$, and $d_t a_t^o \leq 0$.*

[30] When Γ shifts right in the FOSD sense there are two opposite effects that affect the influence of aspirational voters. First, the lowest minimum collateral, $\hat{a}(\text{-}\underline{b}, \Gamma)$, and the maximum occupational threshold, $\tilde{a}(\bar{b}, \Gamma)$, increase (see equations (A.18) and (A.21) in the Appendix). Therefore, the incumbent class—aspirational voters— in Figure 7 encompasses agents that have a weaker preference for business policies (extensive margin effect). Second, since firm profits decrease, the IC constraint becomes more binding for any given level of wealth. Thus, to overcome financial constraints, the poorest agents from the emerging class are willing to support a more pro-business policy (intensive margin effect). In the proof of Proposition 2, I show that the extensive margin effect dominates. Hence, when the wealth distribution shifts right, the fact that the emerging class encompasses agents with weaker preferences for business policies causes an expansion of the welfare state.

(a) *If* $r-\rho<0$ *and* $b_0<0$*, then:* $b^*\in[\text{-}\underline{b},0)$*,* $\tau^*=r-\rho$*, and* $A^*\in(\underline{a},+\infty)$*.*

(b) *If* $b_0>0$*, there are two cases:*

i. $b^*\in[\text{-}\underline{b},\overline{b}]$*,* $\tau^*=r-\rho$*, and* $A^*\in(\underline{a},+\infty)$*.*

ii. $b^*=\text{-}\underline{b}$*,* $\tau^*=\frac{\text{-}\underline{b}\,y(e^{*\prime})(1-e^{*\prime})}{\underline{a}}$*, and* $A^*=\underline{a}$*, where* $e^{*\prime}$ *solves* $h(e^{*\prime})=(1+r)(I-\underline{a})$*.*

The main message of Proposition 2 is that whenever $\tau_0<r-\rho$, social benefits increase over time, the economy accumulates wealth, and the occupational threshold increases. Conversely, if $\tau_0>r-\rho$, these dynamics are reversed. It it important to note that the proposition does not specify the properties of the initial wealth distribution Γ_0 that lead to either case. This is a more complex question that is addressed in Section 6.1.

I start by explaining the evolution of social benefits—the welfare state—which is the main focus of the paper. Consider a discrete-time model where the length of a period is $\Delta>0$. Suppose that Γ_0 is such that $\tau_0<r-\rho$. Then, the initial saving rate out of disposable income is positive ($\theta_0>0$), and thus, agents save a positive fraction of their assets at $t=0$. As a result, the wealth distribution shifts right, i.e. Γ_Δ FOSD Γ_0. The shift in the wealth distribution creates two opposing effects.

First, under the next period wealth distribution (Γ_Δ) and the initial transfer rate (b_0), the mass of entrepreneurs is larger than at $t=0$: $1-\Gamma_\Delta(\hat{a}(b_0,\Gamma_0))>e^*$ (*distributional effect*).

Second, because more agents produce physical capital, the wage rate goes up and firms' profits decrease. As a result, the minimum collateral increases, $\hat{a}(b_0,\Gamma_\Delta)>\hat{a}(b_0,\Gamma_0)$ (*collateral effect*), decreasing to some extent the fraction of entrepreneurs.

In the proof of Proposition 2, I show that the distributional effect dominates. Therefore, under the new distribution Γ_Δ there are "too many" entrepreneurs: $1-\Gamma_\Delta(\hat{a}(b_0,\Gamma_\Delta))>e^*$. Hence, in order to attain the desired level of entrepreneurs e^*, the government must increase social benefits, that is $b_\Delta>b_0$.

The same argument applies for the subsequent periods. As a result, social benefits exhibit an increasing path over time. In the long-run, the spending on social benefits attains its steady state level either when the saving rate becomes zero ($\theta_t = 0$) or when the economy hits its maximum sustainable transfer rate ($b = \bar{b}$). Thus, whenever the initial wealth distribution, Γ_0, is such that $\tau_0 < r - \rho$, social benefits increase over time. The intuition operates in the opposite direction when $\tau_0 > r - \rho$.

Alternatively, these results can be understood by considering dynamics of aspirational voting over time. When $\tau_0 < r - \rho$, the wealth distribution shifts right over time, meaning that, *ceteris paribus*, the fraction of entrepreneurs increases. Thus, as the economy accumulates wealth, firms' profits decrease making it less attractive to start a firm. Thus, the emerging class is less prone to vote for pro-business policies over time (aspirational voting weakens), leading to an increasing path of social benefits.

Figure 8 illustrates in detail Proposition 2. The transition dynamics can be grouped into four general cases. In Cases 1. (a) and 2. (a), all the variables converge smoothly towards a steady state, with the the tax rate converging to $r - \rho$. However, in Cases 1. (b) and 2. (b), the transition dynamics are more involved and two different patterns can emerge within each case. The solid and dashed lines represent scenarios i. and ii. as described in the proposition. Importantly, the difference between both scenarios is the initial wealth distribution, Γ_0.

The figures provide additional insights regarding of the evolution of taxes, which are not included in the proposition for space considerations (e.g. the reversals in Cases 1. (b) and 2. (b)). The proof of the proposition in Appendix A further explores these additional properties. Additionally, in Section C.6 in the Appendix, I illustrate Case 1.(b) by using the OC-IC diagram.

First, in both Cases 1. (b) and 2.(b), there exists an equilibrium in which the economy converges smoothly towards a steady state such that $\tau^* = r - \rho$ (solid lines in Panels 1. (b) and 2. (b)).

Second, in Case 1. (b), the transfer rate b_t reaches its maximum sustainable value $\bar{b}$ before the tax rate can reach its steady state, $r - \rho$ (dotted line in Panel 1. (b)). Once the economy reaches this point, the OC constraint binds, preventing further increases in b_t despite the ongoing wealth accumulation. As a result, the tax rate begins to decrease until it reaches zero. Agents continue accumulating wealth indefinitely and the wealth distribution diverges.

Finally, in Case 2. (b), the economy attains the minimum allowable transfer rate -$\underline{b}$ before τ reaches $r-\rho$ (dotted line in Panel 2. (b)). In this scenario, the wealth distribution collapses to $\underline{a}$, and the tax rate reverses its decreasing trend, converging to some $\tau^* \geq 0$.

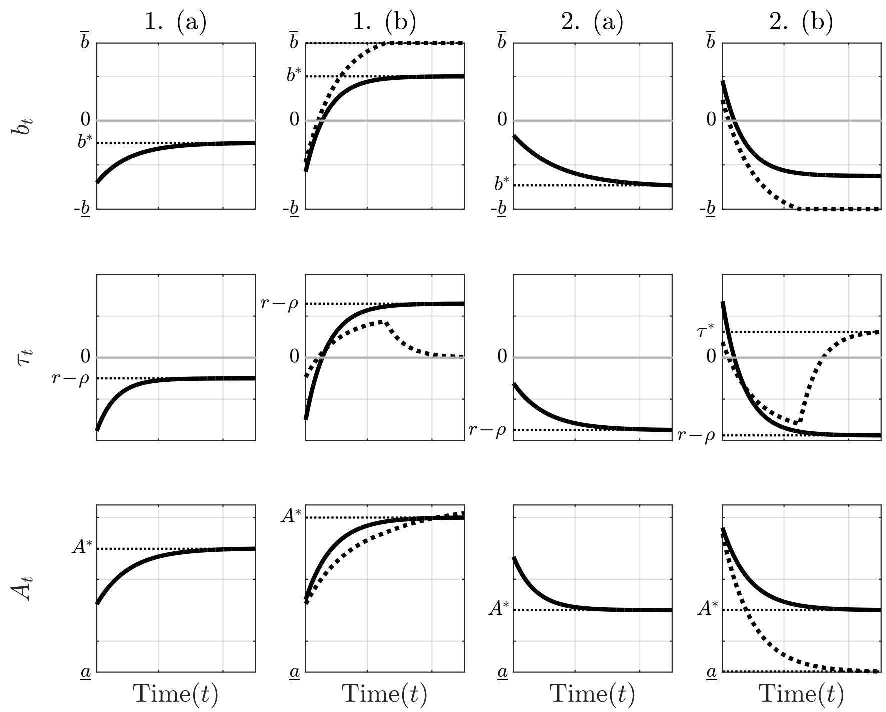


Figure 8: Transition dynamics: Political equilibrium.

In general, the political equilibrium converges to a non-degenerate stationary distribution in the long-run. However, under certain parameter values (Case 1. (b)), the economy continues to accumulate wealth indefinitely, causing the wealth distribution to diverge. In other situations, the wealth distribution collapses to $\underline{a}$ in the long-run (Case 2. (b)). The main difference between each case is the initial wealth distribution, Γ_0.

Regardless of the case, it is possible to provide a theoretical characterization for the evolution of social benefits as measured by b_t, which eventually reaches a stationary level. In many instances, the equilibrium transfer rate is not constrained by the exogenous lower bound -$\underline{b}$, meaning that $b_t \in (-\underline{b}, \overline{b}]$. In certain equilibria, the economy reaches the maximum sustainable transfer rate $\overline{b}$.

6.1 The Initial Wealth Distribution and the Evolution of the Welfare State

The objective of this section is to determine the specific properties of the initial wealth distribution that underlie the different patterns of social benefits studied in the previous section. Given the infinitely dimensional nature of distributions and the limited restrictions imposed on initial distributions so far, obtaining an analytical characterization of such distributional properties is quite challenging. In fact, even quantitatively addressing this question poses significant computational complexity.

To overcome these challenges, I narrow down the set of initial distributions to those generated by applying a Mean Preserving Spread (MPS) on the set of stationary wealth distributions. This approach allows me to classify the initial distributions according to their mean and inequality, and then, characterize the evolution of social benefits depending on these properties. Alternatively, in Section B.6 in the Appendix, I study the transition dynamics when the initial distributions are constructed according to FOSD.

To start with, let $\mathcal{G}_0$ denote the set of initial wealth distributions that are continuously differentiable and with support in $[\underline{a}, +\infty]$. Further, consider the set of stationary distributions $\mathcal{G}^*$ that satisfy equations (3.17) and (3.19). I construct distributions in $\mathcal{G}_0$ by applying an MPS on the distributions in $\mathcal{G}^*$. Given the generated initial distribution Γ_0, I explore the transition dynamics of the equilibrium transfer rate, b_t. This procedure can be interpreted as an MIT shock on the wealth distribution. Figure 9 illustrates the MPS approach.

First, I select the set of stationary distributions $\Gamma^* \in \mathcal{G}^*$ with some mean A^*, represented by the dashed set in Figure 9 and denoted by $\Gamma^*(A^*)$.

Second, I apply an MPS on $\Gamma^*(A^*)$ to construct a set of initial wealth distributions, this step is indicated by the gray arrow in the figure. The MPS set is represented by the dotted set named as $MPS(\Gamma^*(A^*)) \in \mathcal{G}_0$ in the figure. The MPS approach sorts the generated distributions in terms of

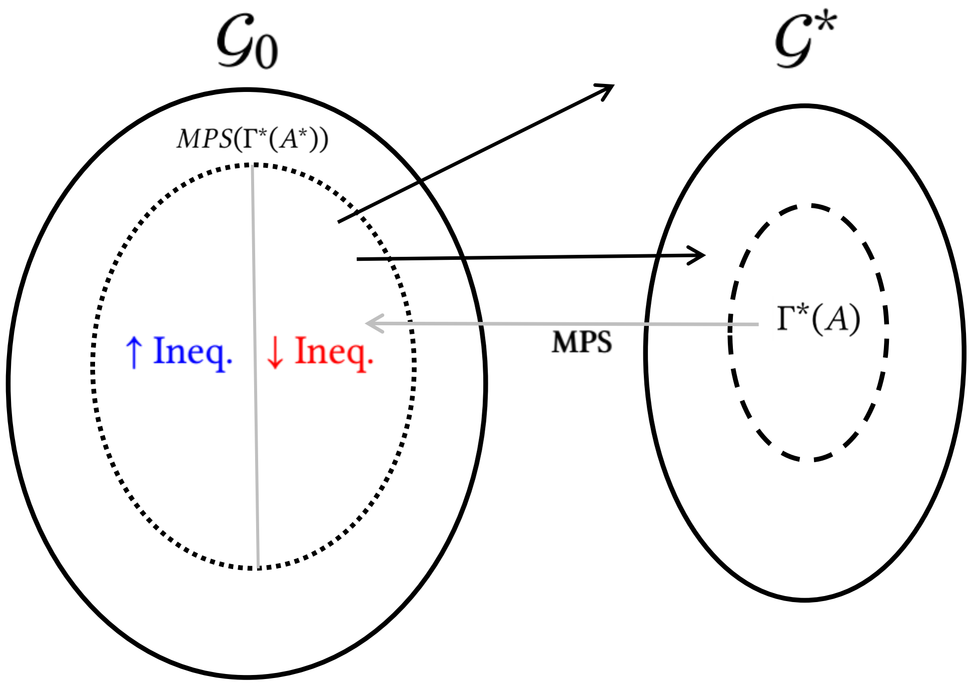


Figure 9: MPS approach.

their inequality relative to $\Gamma^*(A^*)$ (lower and higher inequality).

Finally, given any initial wealth distribution $\Gamma_0 \in MPS(\Gamma^*(A^*))$, I study the transition dynamics. In general, the economy converges to a new stationary wealth distribution, as indicated by the horizontal black arrow in the figure. In some cases, the wealth distribution may not even attain a steady state in the long-run (see Case 1. (b) in Figure 8), as indicated by the black arrow that falls outside $\mathcal{G}^*$. However, I can still characterize the evolution of the size of the welfare state that always attains a stationary level.[31]

Once this process is completed, we can go in the other direction and ask: What are the transition dynamics given some initial distribution Γ_0? To answer this question, pick some initial distribution $\Gamma_0 \in \mathcal{G}_0$ with some mean A. If the initial distribution belongs to the dotted set in Figure 9, then the transition dynamics can be characterized in terms of its aggregate wealth and inequality. If not, then in general the direction of social benefits over time is analytically ambigu-

[31]Note that the MPS approach is not equivalent to studying a standard impulse response function. First, given the new wealth distribution Γ_0, the transfer rate may reach a new stationary wealth distribution in the long-run. In general, after an MPS shock, the economy transitions between two steady states. Second, the shock that hits the economy is infinite dimensional as it shifts the entire wealth distribution.

ous. This motivates the quantitative exercise in Section 7.

6.1.1 Mean Preserving Spread (MPS)

I restrict attention to MPS distributions that intersect only once at some asset level $\tilde{A}$ that belongs to some neighbourhood around the mean A of the distribution (denoted as $\tilde{A} \in N(A)$). The seminal paper by Rothschild and Stiglitz (1971) used a similar approach to study the economic consequences of risk. More recently, Fischer and Huerta (2021) study the impact of the same kind of MPS distributions on financial and labor policies in a static model with heterogeneous agents.

Definition 1 *Consider two distributions,* Γ_1 *and* Γ_2*, with mean* A *and support in* $[\underline{a}, +\infty)$. Γ_2 *is said to be an MPS of* Γ_1*, denoted as* $\Gamma_2 >_{MPS} \Gamma_1$*, if:*

1. $\Gamma_2(a) > \Gamma_1(a)$ *if* $a < \tilde{A}$.

2. $\Gamma_2(a) \leq \Gamma_1(a)$ *if* $a > \tilde{A}$.

Further, the intersect of both distributions satisfies $\tilde{A} \in N(A)$.

The advantage of using MPSs is that I can isolate the pure impact of higher inequality on the evolution of the equilibrium transfer rate. An MPS is equivalent to second order stochastic dominance while keeping the mean unchanged. Shorrocks (1983) shows that second order stochastic dominance is equivalent to an ordering according to the generalized Lorentz curve. In fact, an MPS is analogous to Lorentz dominance. Thus, the results can be interpreted in terms of standard inequality measures. In particular, the generated initial distributions Γ_0 can be ordered in terms of their mean and inequality relative to some stationary distribution Γ^*.

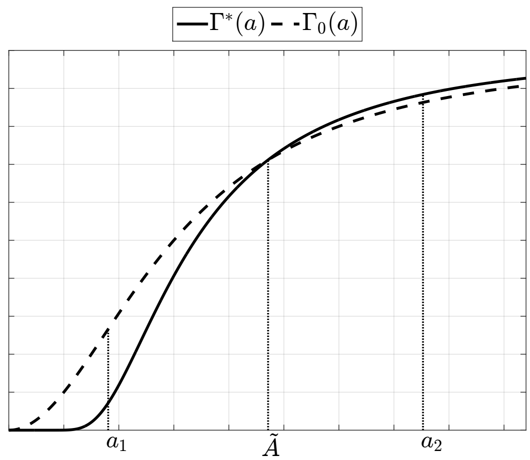
In order to state the next proposition, define the set of initial distributions that are an MPS of the stationary distribution $\Gamma^* \in \mathcal{G}$ as:

$$\mathcal{G}_{MPS}(\Gamma^*) \equiv \{\Gamma_0 \, : \, \Gamma_0 >_{MPS} \Gamma^*, \Gamma^* \in \mathcal{G}\}.$$

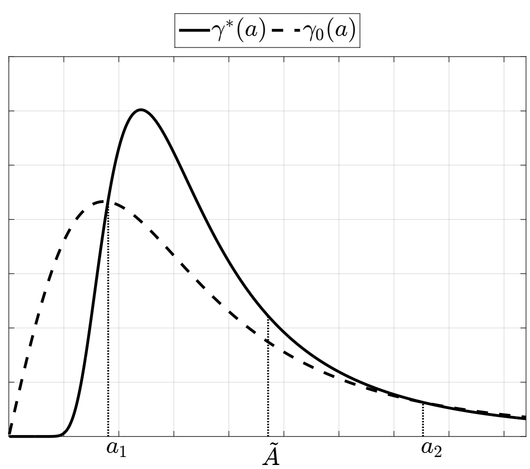
Note that $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$ is equivalent to say that Γ^* Lorentz dominates Γ_0. That is, the Lorentz curve of Γ^* lies everywhere above that of Γ_0. Hence , Γ_0 is more unequal than Γ^*.

Analogously, I define the set of reverse-mean preserving spreads (reverse-MPS) of some steady-state distribution $\Gamma^* \in \mathcal{G}$ as follows:

$$\mathcal{G}_{MPS_{rev}}(\Gamma^*) \equiv \{\Gamma_0 \,:\, \Gamma_0 <_{MPS} \Gamma^*, \Gamma^* \in \mathcal{G}\}.$$



(a) Cumulative distribution functions, Γ^* and Γ_0



(b) Density functions, γ^* and γ_0.

Figure 10: MPS distributions, $\Gamma^* \in \mathcal{G}$ and $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$

Figure 10 illustrates the properties of the MPSs curves in Definition 1. Figure 10a depicts the cumulative distribution for $\Gamma^* \in \mathcal{G}$ and $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$, i.e. when Γ_0 (dashed line) is an MPS of Γ^* (solid-line). As shown in the figure, the cumulative distributions intersect in the interior only once at $\tilde{A}$ (single-crossing property) . Figure 10b shows the density functions which intersect in the interior only twice (double-crossing) at some wealth levels a_1 and a_2, with $a_1 < a_2$. The first cross (a_1) corresponds to the wealth level at which $\Gamma_0 - \Gamma^*$ is maximized, while the second cross (a_2) minimizes this difference. Graphically, the MPS distribution Γ_0 shifts the frequencies of Γ^* from the middle towards the tails, while keeping the mean unchanged.

6.1.2 Main Result: The Predicted Evolution of the Welfare State

The following proposition describes the evolution of the welfare state when the initial wealth distributions are ordered according to the MPS approach:

Proposition 3 *Consider some stationary wealth distribution* $\Gamma^* \in \mathcal{G}$ *with mean* A. *The initial wealth distribution is such that* $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$, *with occupational threshold* $a_0^o = a^o(\Gamma_0)$. *Denote by* $\tilde{A} \in N(A)$ *the intersect of both cumulative distributions. Then,*

1. *If* $\tilde{A} < a_0^o$*: i)* $d_t b_t \leq 0$, *ii)* $d_t A_t \leq 0$, *iii)* $d_t a_t^o \leq 0$.

2. *If* $\tilde{A} > a_0^o$, *the signs of i) to iii) are reversed.*

The transition dynamics are reversed if $\Gamma_0 \in \mathcal{G}_{MPS_{rev}}(\Gamma^*)$.

Proposition 3 characterizes the transition dynamics of the political equilibrium as a function of initial inequality and wealth. The Cases 1. and 2. compare the mean (A) and the minimum wealth level to start a firm (a_0^o). Countries that initially satisfy $A < a_0^o$ can be interpreted as *poor* (or low wealth) because the average agent does not have sufficient wealth to start a firm. In contrast, countries in which $A > a_0^o$ are said to be *rich* (or high wealth).

Proposition 3 reads as follows. 1) Countries that start poor and and unequal, exhibit a decreasing path of spending on social benefits over time . 2) Countries that are initially rich and unequal have increasing social benefits over time. Items 1) and 2) apply in the opposite direction when countries start with a more equal distribution, i.e. when $G_0 \in \mathcal{G}_{MPS_{rev}}(G)$. Table 1 summarizes these results.

	Unequal	Equal
Poor ($A_0 < a_0^o$)	$\searrow b$	$\nearrow b$
Rich ($A_0 > a_0^o$)	$\nearrow b$	$\searrow b$
	United States (1970-2019)	*Sweden (1995-2019)*

Table 1: Initial distribution and the evolution of social benefits

6.1.3 The American and Swedish Experience

The results presented in Table 1 are qualitatively consistent with some clear trends observed in the data, as discussed in Section D.1 in the Appendix. For instance, spending on social benefits has been consistently increasing in the US since 1970. On the other hand, perhaps surprisingly, social benefits have been decreasing in Sweden since 1995. These trends are robust to using different measures, such as cash transfers only, incorporating in-kind social benefits, accounting for tax breaks, or using post-tax social benefits.

According to the World Inequality Database (WID)), inequality in the US in 1970 was high relative to the rest of the world (Gini index exceeding 0.8). On the other hand, Sweden had relative low inequality in 1995 (Gini index around 0.7). Both countries are considered high wealth, and thus, their experiences align with the last row of Table 1.

In what follows, I provide intuition for the American experience. First, since the country starts with high wealth but unequal there is a small incumbent class that has enough wealth to start a firm. Thus, initially there are “few" firms, low competition, and thus, high profits. Since the economy starts rich, there is a large emerging class that demands pro-business policies (i.e. low social benefits) hoping to be able to overcome financial constraints and start a firm (aspirational voting is strong). As a result, social benefits are initially low.

As time goes by, many agents from the emerging class benefit from pro-business policies and can start their businesses. Thus, profits decrease, reducing the attractiveness of entrepreneurship. As a result, the emerging class joins the working class in demanding higher social benefits (i.e. aspirational voting weakens).

In the long-run, the economy becomes wealthier, and thus, the government can finance more social benefits for workers by applying a relatively low tax to businesses. As a result, having higher social benefits becomes politically sustainable in the future. The intuition is reversed for a country as Sweden that starts with high wealth but more equally distributed.

It is important to emphasize that the qualitative examples discussed above are aimed just to illustrate the underlying mechanisms of the model. In the next section, I perform a quantitative exercise in a set of countries to assess the model's ability to predict the evolution of social benefits over time.

7 Quantitative Analysis

In this section, I test whether the model can predict the observed trends of social benefits in 24 selected countries from all continents. These countries include: Australia, Belgium, Canada, Denmark, Spain, Finland, France, Great Britain, Germany, Greece, Hungary, Ireland, Israel, Italy, South Korea, Latvia, Luxembourg, Norway, New Zealand, Poland, Portugal, Sweden, United States, and South Africa.

Guided by the theoretical results, I calibrate the model while taking as given the observed wealth distribution of each country at a starting year. Then, I simulate the model for the subsequent years to test whether the generated path of social benefits is consistent with that of the data. Due to data limitations, for most of the countries, the time-horizon considered is from 1995 to 2019. Thus, given the empirical wealth distribution of each country in 1995, I test whether the model can predict the dynamics of social benefits for the next 25 years.[32]

[32] In the case of the United States, the data starts in 1970, and for France, it starts from 1980. Thus, I assess whether the model can predict the trend for the next fifty and forty years, respectively.

Overall, the model predicts the trend of social benefits in 18 out of 24 evaluated countries (75% prediction rate). Moreover, in some cases, the model can predict not only the sign of the trend but also its magnitude and shape. Notable examples are Canada, Sweden, and the US. The quantitative exercise provides strong support for the conclusion that the wealth distribution is a key force behind the striking differences in the evolution of the welfare state across countries.

In Subsections 7.1 to 7.3, I provide a brief description of the quantitative exercise. For space considerations, in Subsection 7.4, I provide the predicted paths of social benefits for four countries: Canada, the United States, Norway, and Sweden. I also interpret the American experience through the lens of the model. In Section E in the Appendix, I provide a detailed description of the calibration and simulation procedures and present the simulated trends for all the countries under consideration.

7.1 Quantitative model

For computational purposes, the model is written in discrete time where the length of a period is Δ, with Δ small. The starting year of the simulation is denoted by T_0, the final year is $T = 2019$.[33]

The economy is the same as the baseline model of Section 2 except for two differences. First, I use a constant risk aversion utility function (CRRA): $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with a standard relative risk aversion coefficient $\sigma = 2$. Second, I introduce a new parameter that captures the "government responsiveness" to aggregate productivity shocks: $\omega \in [0, 1]$. At any period $t+\Delta$, the government chooses the transfer rate $b_{t+\Delta}$ that satisfies:

$$e_{t+\Delta} = (1-\omega) \cdot e^*_{t+\Delta} + \omega \cdot e_t, \tag{7.1}$$

where $e^*_{t+\Delta}$ is the optimal fraction of entrepreneurs given $(\Gamma_{t+\Delta}, Z_{t+\Delta})$ and e_t is the fraction of entrepreneurs of the previous period. This specification helps to smooth the transfer rate responsiveness to aggregate changes in productivity along the transition path. In particular, it helps

[33] For the United States $T_0 = 1970$, while for France $T_0 = 1980$. For the rest of the countries, $T_0 = 1995$.

to maintain the stability of the model at T_0, that is, that in the neighborhood of T_0 the economy does not experience drastic deviations from its initial state. The parameter ω admits at least three interpretations.

First, it can be interpreted as capturing country-specific political constraints that limit the rate at which the government can adjust social benefits to its desired level. This can happen due to previous political compromises or domestic political opposition. When $\omega = 0$, the government can freely choose the transfer rate to its desired level. However, when $\omega \to 1$ the government faces serious political constraints.

Second, in the model social benefits are adjusted continuously. However, in reality government spending is decided at a much lower frequency. Thus, ω can also account for country-specific rates of adjustment of social spending.

Finally, ω can capture government spending leakages, meaning that not all allocated funds translate into social benefits for the intended recipients due to corruption or inefficiencies. Related to this idea, ω may also reflect inefficiencies in the tax system. Even when the government may want to increase social benefits, it may not be able to raise the necessary funds because it cannot collect taxes efficiently.

7.2 Calibration

The quantitative exercise requires choosing the model's parameters for 24 countries. The main objective of the calibration is to match the level of social benefits of each country given the empirical wealth distribution in 1995. I also match other data moments such as the capital to labor ratio, the investment to output ratio, and the Gini income coefficient. I choose seven country-specific parameters, among the most important ones are: the interest rate r, the discount factor ρ, and the political weight ϕ. I select the country-specific parameters by using the Generalized Method of Moments. Section E.2 in the Appendix provides more details about the calibration strategy.

7.3 Simulation

The model takes as input two country-specific measures. First, the wealth distribution at a starting year, Γ_0, which depends on the earliest available observation for each country. Second, the production function coefficient α plus the aggregate productivity path $\{Z_t\}_{T_0}^{T}$ of each country estimated via Solow residuals. To obtain the simulated path of social benefits, $\{b_t\}_{T_0}^{T}$, I iterate forward the KF equation (3.15) and the political equilibrium condition (5.4) taking as given Γ_0 and $\{Z_t\}_{T_0}^{T}$. In Section E.3 in the Appendix, I describe in detail the simulation procedure.

7.4 Quantitative results

Figure 11 presents the evolution of social benefits as a share of GDP for the four representative countries considered in the introduction: Canada, United States, Norway, and Sweden. The black lines correspond to the data, while the gray lines to the simulated social benefits. The dotted lines represent the trends. The starting year corresponds to the earliest year from which a country has data available for all the variables used in the simulation. Figure 21 in Section E.4 in the Appendix compares the trend of social benefits from the data and the model for the 24 evaluated countries.

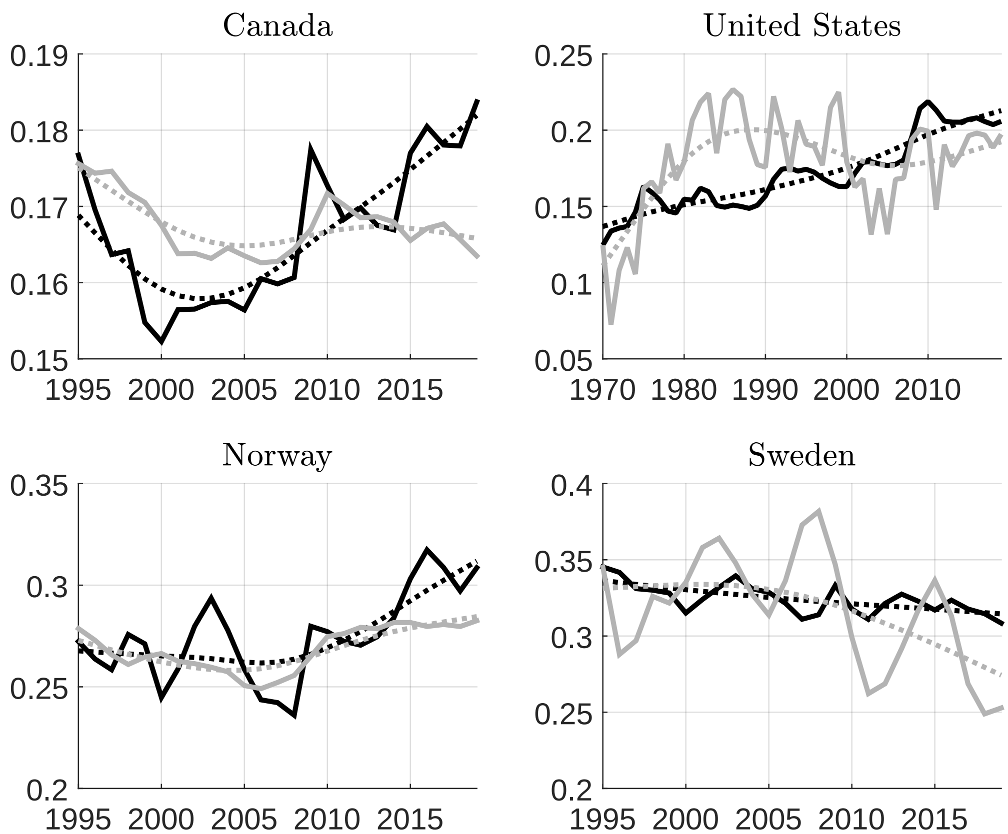


Figure 11: Social benefits (% GDP). Black: data, Gray: model, Dotted line: trend.

Figure 11 shows that the model is able to predict not only the sign of the trend of social benefits but also in some cases the magnitude and shape of the trend. First, the model successfully predicts the increasing path observed in the US in the last fifty years based on the wealth distribution in 1970. Second, the model is able to predict the "U-shape trend" for Canada between 1995 and 2010. Finally, among Scandinavian countries, the model captures the initially somewhat flat trend in Norway and then predicts the increasing trend starting around 2008. The model also forecasts a decreasing trend in Sweden since 1995 as in the data.

7.4.1 The American experience

In what follows, I interpret the American experience through the lens of the model. The Panel a) of Figure 12 shows the trends of social benefits as a fraction of GDP. The solid line corresponds to the data. The dotted line is the predicted trend by the model when it incorporates the estimated aggregate productivity presented in Panel b) (*Full Model*). The dashed line is the prediction of the model when aggregate productivity remains fixed over time (*Base Model*).

The Panel c) and Panel d) in Figure 12 show the evolution of the share of population that belongs to the middle and lower class, respectively. The solid line presents the estimated shares based on the analysis of the Pew Research Center (2022) using the Current Population Survey (CPS). The dashed and dotted line depict the predicted shares by the model. In the model, the middle class is the fraction of agents that belong to the emerging class in Figure 7 in Section 5.2. The lower class corresponds to the working class.[34]

The Base Model The dashed line in Figure 12 isolates the model from exogenous changes in productivity. Thus, the generated path is solely a result of the endogenous feedback between the wealth distribution and the size of the welfare state. The model predicts an increasing path as in the data just based on the observed wealth distribution in 1970. In particular, the US started in the 70's with sufficiently high inequality such that initially a relatively small incumbent class had enough wealth to start a firm (less than 20% according to the dashed lines in Panels c) and d)). Thus, the intuition for why social benefits exhibit an increasing path over time works as the one provided for Table 1 in Section 6.1.

Full Model To understand the results of the full model (dotted line in Panel a)), in Section C.8 in the Appendix, I study the impact of a permanent increase in aggregate productivity. Intuitively, higher productivity increases firms' profits and relaxes financial constraints. From the point of view of the middle wealth agents, it becomes more attractive to start a firm while at the same

[34] In the quantitative exercise I calibrate the minimum transfer rate, -$\underline{b}$, to match the fraction of agents in the middle class in 1970. The value of -$\underline{b}$ is never binding in equilibrium, thus it does not constraint the predicted path of social benefits.

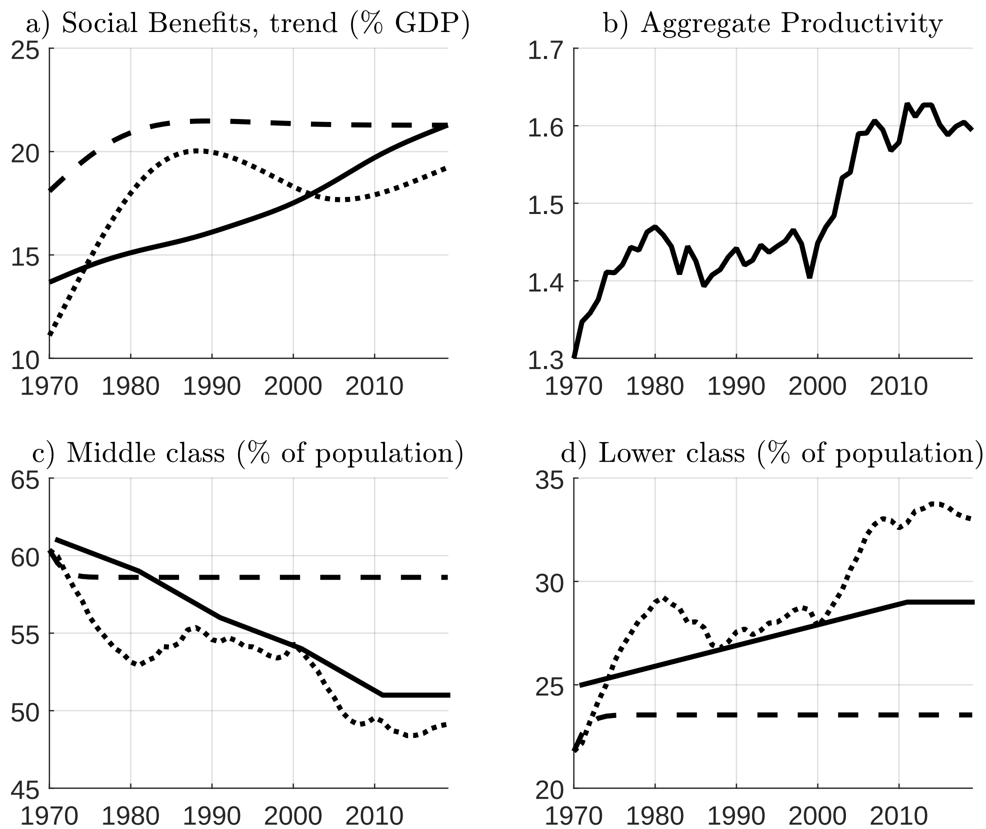


Figure 12: Solid line: Data, Dashed line: Base Model, Dotted line: Full Model.

time it is easier to obtain finance. Thus, a positive productivity shock expands the emerging class, and thus, strengthen aspirational voting, decreasing social benefits. Then, why have been social benefits increasing in the last fifty years in the US despite rising productivity?

The answer to that question depends on the general equilibrium effects that follow just right after an increase in productivity. The decrease in social benefits due to higher productivity fosters wealth accumulation, and thus, increases the next period mass of agents that can start a firm. The wealthier agents from the emerging class join the incumbent class and become entrepreneurs. However, higher competition decreases firm profits tightening collateral constraints.

From the point of view of poorest agents from the middle class, entrepreneurship becomes less attractive and it becomes harder to receive finance to start a business. As a result, the poorest agents from the emerging class join the working class in demanding higher social benefits. As shown by the dotted line in Panel c) and d), the middle class shrinks over time, while the working class expands. These changes in the composition of the population translate into a decreased demand for pro-business policies and an increased demand for social benefits over time, explaining the increasing path of social benefits observed in the US in the last fifty years.

Overall, the increasing path of social benefits in the US in the last fifty years responds to the the shrinking middle class and expanding lower class. These features are consistent with the data as can be seen by comparing the dotted and black lines in Panels c) and d) in Figure 12. The model predicts these changes in the composition of both classes based on the wealth distribution in 1970. The model correctly predicts the increasing path of social benefits over time based on the endogenous preferences for social benefits of each class.

7.4.2 Counterfactual Analysis

What is the role of changes in the political orientation of the government in explaining the evolution of the welfare state?

To address this question, I perform a counterfactual analysis in three countries that exhibit diametrically different trends in social benefits: Canada, the US, and Sweden. The objective of this exercise is to evaluate the importance of changes in the political orientation of the government for the trend of social benefits. The exercise works as follows.

First, for each of these countries, I solve for the path of political weights ($\{\phi_t\}_{T_0}^{T}$) that matches the observed path of social benefits. Second, I construct two extreme scenarios around $\{\phi_t\}_{T_0}^{T}$: a "highly pro-worker" and a "highly pro-business" scenario. Finally, I simulate the evolution of social benefits and verify whether the trend of social benefits would have changed under these two scenarios. Section E.5 in the Appendix describes the exercise and presents detailed results for the US.

The main finding from this counterfactual exercise is that the trend of social benefits would not have changed significantly in either country. Thus, there is a limited role of changes in the government in explaining the evolution of the welfare state. In conclusion, the counterfactual exercise confirms the key role of the wealth distribution in explaining the dynamics of the welfare state.

8 Conclusions

This paper develops a theory that explains the striking differences in the evolution of the welfare state across countries in the last few decades. The model is structured around a single force: the dynamic interaction between the wealth distribution and social benefits over time. According to the model, the spending on social benefits is expected to increase in rich countries with high wealth inequality, while it should decline if inequality is low. The model highlights the crucial role of an endogenous middle class that is willing to sacrifice social benefits and support pro-business policies aspiring to become entrepreneurs. A calibrated version of the model successfully predicts the observed trends of spending on social benefits in 18 out of 24 countries from all continents. The main conclusion is that the wealth distribution is a key force behind the striking differences in the evolution of the welfare state across countries.

This article paves the road for the development of theories to understand the dynamic interaction between policies, inequality, and macroeconomic variables over time. I have focused on studying how the dynamic interaction between the wealth distribution and policies—*the inequality-policy link*—influences the evolution of the welfare state. However, the inequality-policy link can also have important consequences for growth. In particular, in my model, depending on the initial wealth distribution some countries can overcome financial constraints in the long-run, while others never escape from financial constraints. This feature is not longer true when policies are exogenously given over time. In that case, countries can always overcome financial constraints in the long-run. This suggests that the inequality-policy link may play an important role when answering a traditional question in economics: Why some countries develop and why others remain underdeveloped?

The tractability of my model allows to address this question through the lens of a theory that incorporates an inequality-policy link over time. Specifically, it can allow for a characterization of the primitives that explain why some countries develop over time and others do not.

References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin, "Social Mobility and Stability of Democracy: Reevaluating de Tocqueville," *The Quarterly Journal of Economics*, 2018, *133* (2), 1041–1105.

_ , **Michael Golosov, and Aleh Tsyvinski**, "Political Economy of Mechanisms," *Econometrica*, 2008, *76* (3), 619–641.

_ , **Mikhail Golosov, and Aleh Tsyvinski**, "Dynamic Mirrlees Taxation under Political Economy Constraints," *The Review of Economic Studies*, 2010, *77* (3), 841–881.

_ , _ , **and** _ , "Political Economy of Ramsey Taxation," *Journal of Public Economics*, 2011, *95* (7-8), 467–475.

Acharya, Sushant, Edouard Challe, and Keshav Dogra, "Optimal Monetary Policy according to HANK," *American Economic Review*, 2023, *113* (7), 1741–1782.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 2022, *89* (1), 45–86.

Alesina, Alberto and Dani Rodrik, "Distributive Politics and Economic Growth," *The Quarterly Journal of Economics*, 1994, *109* (2), 465–490.

_ **and Eliana La Ferrara**, "Preferences for Redistribution in the Land of Opportunities," *Journal of Public Economics*, 2005, *89* (5-6), 897–931.

_ **and George-Marios Angeletos**, "Fairness and Redistribution," *American Economic Review*, 2005, *95* (4), 960–980.

Audretsch, David B, Maksim Belitski, Farzana Chowdhury, and Sameeksha Desai, "Necessity or Opportunity? Government Size, Tax Policy, Corruption, and Implications for Entrepreneurship," *Small Business Economics*, 2022, *58* (4), 2025–2042.

Bachmann, Rüdiger and Jinhui H Bai, "Politico-Economic Inequality and the Comovement of Government Purchases," *Review of Economic Dynamics*, 2013, *16* (4), 565–580.

Banerjee, Abhijit and Andrew Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, *101* (2), 274–298.

Benabou, Roland and Efe A Ok, "Social Mobility and the Demand for Redistribution: the POUM Hypothesis," *The Quarterly Journal of Economics*, 2001, *116* (2), 447–487.

— **and Jean Tirole**, "Belief in a Just World and Redistributive Politics," *The Quarterly Journal of Economics*, 2006, *121* (2), 699–746.

Bernanke, Ben and Mark Gertler, "Agency Costs, Net Worth, and Business Fluctuations," *The American Economic Review*, 1989, *79* (1), 14–31.

Buera, Francisco J and Benjamin Moll, "Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity," *American Economic Journal: Macroeconomics*, 2015, *7* (3), 1–42.

— **and Yongseok Shin**, "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

— **, Joseph P Kaboski, and Yongseok Shin**, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 2011, *101* (5), 1964–2002.

Cagetti, Marco and Mariacristina De Nardi, "Entrepreneurship, frictions, and wealth," *Journal of Political Economy*, 2006, *114* (5), 835–870.

Farhi, Emmanuel, Christopher Sleet, Ivan Werning, and Sevin Yeltekin, "Non-Linear Capital Taxation Without Commitment," *Review of Economic Studies*, 2012, *79* (4), 1469–1493.

Fischer, Ronald and Diego Huerta, "Wealth Inequality and the Political Economy of Financial and Labour Regulations," *Journal of Public Economics*, 2021, *204*, 104553.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *The Review of Economic Studies*, 1993, *60* (1), 35–52.

Gonzalez-Eiras, Martin and Dirk Niepelt, "The Future of Social Security," *Journal of Monetary Economics*, 2008, *55* (2), 197–218.

Hassler, John, José V Rodríguez Mora, Kjetil Storesletten, and Fabrizio Zilibotti, "The Survival of the Welfare State," *American Economic Review*, 2003, *93* (1), 87–112.

_ , **Kjetil Storesletten, and Fabrizio Zilibotti**, "Dynamic Political Choice in Macroeconomics," *Journal of the European Economic Association*, 2003, *1* (2-3), 543–552.

_ , **Per Krusell, Kjetil Storesletten, and Fabrizio Zilibotti**, "The Dynamics of Government," *Journal of Monetary Economics*, 2005, *52* (7), 1331–1358.

Henrekson, Magnus, "Entrepreneurship: a weak link in the welfare state?," *Industrial and Corporate Change*, 2005, *14* (3), 437–467.

Hessels, Jolanda, André van Stel, Peter Brouwer, and Sander Wennekers, "Social Security Arrangements and Early-Stage Entrepreneurial Activity," *Comp. Lab. L. & Pol'y J.*, 2006, *28*, 743.

_ , **Marco Van Gelderen, and Roy Thurik**, "Drivers of Entrepreneurial Aspirations at the Country Level: the Role of Start-Up Motivations and Social Security," *International Entrepreneurship and Management Journal*, 2008, *4*, 401–417.

Huerta, Diego, "The Political Economy of Labor Policy," 2023.

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies with Financial Frictions," *Econometrica*, 2019, *87* (1), 139–173.

Jang, Youngsoo, "Time-Consistent Taxation with Heterogeneous Agents," 2023.

Koellinger, Philipp and Maria Minniti, "Unemployment Benefits Crowd out Nascent Entrepreneurial Activity," *Economics Letters*, 2009, *103* (2), 96–98.

Krusell, Per and Jose-Victor Rios-Rull, "Vested Interests in a Positive Theory of Stagnation and Growth," *The Review of Economic Studies*, 1996, *63* (2), 301–329.

— **and** — , "On the Size of US Government: Political Economy in the Neoclassical Growth Model," *American Economic Review*, 1999, *89* (5), 1156–1181.

— , **Vincenzo Quadrini, and Jose-Victor Rios-Rull**, "Are Consumption Taxes Really Better than Income Taxes?," *Journal of Monetary Economics*, 1996, *37* (3), 475–503.

— , — , **and José-Víctor Ríos-Rull**, "Politico-Economic Equilibrium and Economic Growth," *Journal of Economic Dynamics and Control*, 1997, *21* (1), 243–272.

Leventoğlu, Bahar, "Social Mobility and Political Transitions," *Journal of Theoretical Politics*, 2005, *17* (4), 465–496.

— , "Social Mobility, Middle Class, and Political Transitions," *Journal of Conflict Resolution*, 2014, *58* (5), 825–864.

Lindbeck, Assar and Jörgen W Weibull, "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice*, 1987, *52* (3), 273–297.

Matsuyama, Kiminori, "Financial Market Globalization, Symmetry-Breaking, and Endogenous Inequality of Nations," *Econometrica*, 2004, *72* (3), 853–884.

Meltzer, Allan H and Scott F Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, *89* (5), 914–927.

Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 2014, *104* (10), 3186–3221.

Nuño, Galo and Benjamin Moll, "Social Optima in Economies with Heterogeneous Agents," *Review of Economic Dynamics*, 2018, *28*, 150–180.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Pecoraro, Brandon, "Why Don't Voters 'Put The Gini Back in the Bottle'? Inequality and Economic Preferences for Redistribution," *European Economic Review*, 2017, *93*, 152–172.

Persson, Torsten and Guido Tabellini, "Is Inequality Harmful for Growth?," *The American Economic Review*, 1994, *84* (3), 600–621.

_ and _ , *Political Economics: Explaining Economic Policy*, The MIT Press, 2000.

Piazzesi, Monika and Martin Schneider, "Housing and Macroeconomics," *Handbook of Macroeconomics*, 2016, *2*, 1547–1640.

Piketty, Thomas, "Social Mobility and Redistributive Politics," *The Quarterly Journal of Economics*, 1995, *110* (3), 551–584.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, "Investor Protection and Corporate Governance," *Journal of Financial Economics*, 2000, *58* (1-2), 3–27.

Quadrini, Vincenzo and JV Rios-Rull, "International Tax Competition with Rising Intangible Capital and Financial Globalization," Technical Report, mimeo 2023.

Rajan, Raghuram G and Luigi Zingales, "The Great Reversals: The Politics of Financial Development in the Twentieth Century," *Journal of Financial Economics*, 2003, *69* (1), 5–50.

_ and Rodney Ramcharan, "Land and Credit: A Study of the Political Economy of Banking in the United States in the Early 20th Century," *The Journal of Finance*, 2011, *66* (6), 1895–1931.

Rothschild, Michael and Joseph E Stiglitz, "Increasing Risk II: Its Economic Consequences," *Journal of Economic Theory*, 1971, *3* (1), 66–84.

Sargent, Thomas J, *The Conquest of American Inflation*, Princeton University Press, 1999.

Shleifer, Andrei and Daniel Wolfenzon, "Investor Protection and Equity Markets," *Journal of Financial Economics*, 2002, *66* (1), 3–27.

Shorrocks, Anthony F., "Ranking Income Distributions," *Economica*, 1983, *50* (197), 3–17.

Sleet, Christopher and Şevin Yeltekin, "Politically Credible Social Insurance," *Journal of Monetary Economics*, 2008, *55* (1), 129–151.

Solomon, Shelby J, Joshua S Bendickson, Matt R Marvel, William C McDowell, and Raj Mahto, "Agency Theory and Entrepreneurship: A Cross-Country Analysis," *Journal of Business Research*, 2021, *122*, 466–476.

Solomon, Shelby, Joshua S Bendickson, Eric W Liguori, and Matthew R Marvel, "The Effects of Social Spending on Entrepreneurship in Developed Nations," *Small Business Economics*, 2022, pp. 1–13.

Song, Changhyeon, Kyung Min Park, and Yeonbae Kim, "Socio-Cultural Factors Explaining Technology-Based Entrepreneurial Activity: Direct and Indirect Role of Social Security," *Technology in Society*, 2020, *61*, 101246.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 2012, *80* (6), 2785–2803.

Stöllinger, Roman and Mario Holzner, "State aid and Export Competitiveness in the EU," *Journal of Industry, Competition and Trade*, 2017, *17*, 203–236.

Tocqueville, Alexis De, *De la Democratie en Amerique*, Hauman, 1835.

Yared, Pierre, "Politicians, Taxes and Debt," *The Review of Economic Studies*, 2010, *77* (2), 806–840.

A Appendix: Main Proofs

A.1 Proof of Lemma 1

Lemma 1 *The optimal consumption and savings policy functions are linear functions of disposable income $y_t(a)$:*

$$c_t(a) = (1-\theta_t)\cdot y_t(a),$$
$$s_t(a) = \theta_t \cdot y_t(a),$$

where the saving rate is: $\theta_t = \left(1 - \frac{\rho}{r-\tau_t}\right)$.

Proof: In order to obtain the policy functions $c(a)$ and $s(a)$, guess that the value function is given by:

$$v_t(a) = B\log(y(a)) + B\tilde{v}_t. \tag{A.1}$$

Rewrite the HJB equation as follows:

$$\rho v_t(a) = \max_c \left\{ \log(c) + \frac{d}{dt}(v_t(a)) \right\}.$$

Using the guess (A.1) and the assumption that agents do not predict the future joint dynamics of (b_t, Γ_t), the second term in the HJB equation is:

$$\frac{d}{dt}(v_t(a)) = B\frac{1}{y(a)}\left[(r-\tau)\right]\dot{a} + Bd_t\tilde{v}_t$$

The budget constraint is: $\dot{a} = y(a) - c$, thus $\frac{d}{dt}(v(a)) = B(r-\tau)\left(1 - \frac{c}{y(a)}\right) + Bd_t\tilde{v}_t$. Then, the HJB equation reads as:

$$\rho v_t(a) = \max_c \left\{ \log(c) + B(r-\tau)\left(1-\frac{c}{y}\right) + Bd_t\tilde{v}_t \right\}.$$

Take the first order condition to obtain: $c = \frac{y(a)}{B(r-\tau)}$. Replacing in the HJB equation gives

$$\rho B \log(y(a)) + \rho B \tilde{v}_t = \log(y(a)) - \log(B(r-\tau)) + B(r-\tau) - 1 + B d_t \tilde{v}_t \tag{A.2}$$

Collecting terms involving $\log(y(a))$ gives that $B = \frac{1}{\rho}$. Therefore, the consumption policy function is $c(a) = \frac{\rho}{r-\tau} \cdot y(a)$. Then, the individual budget condition gives $\dot{a} = s(a) = \left(1 - \frac{\rho}{r-\tau}\right) \cdot y(a)$ as stated in the lemma. Finally, the value function is $v_t(a) = \frac{1}{\rho}\left(\log(y) + \tilde{v}_t\right)$ where

$$\begin{aligned} \rho \tilde{v}_t &= \rho \log\left(\frac{\rho}{r-\tau}\right) + r - \tau - \rho + d_t \tilde{v}_t, \\ &= \rho \log(1-\theta_t) + \frac{\rho\theta_t}{1-\theta_t} + d_t \tilde{v}_t. \end{aligned}$$

■

A.2 Proof of Lemma 2

Lemma 2 *The evolution of the cumulative wealth distribution $\Gamma_t(a)$ is characterized by the Kolmogorov Forward (KF) equation:*

$$d_t \Gamma_t(a) = -\Gamma_t(a_t^o) \cdot s_t(a) d_a \Gamma_t(a) - \left(1 - \Gamma_t(a_t^o)\right) \cdot s_t(a) d_a \Gamma_t(a).$$

Proof: Assets evolve according to:

$$da = s(a)dt, \tag{A.3}$$

where $s(a)$ is given by equation (3.8). Consider now a discrete time model where the length of a period is Δ. The analogue of (A.3) is:

$$a_{t+\Delta} - a_t = \int_t^{t+\Delta} s(a_h) dh.$$

For Δ small, $\int_t^{t+\Delta} s(a_h)dh \approx \Delta s(a_t)$ and also $\int_t^{t+\Delta} s(a_h)dh \approx \Delta s(a_{t+\Delta})$. Using the second approximation and rearranging terms:

$$a_{t+\Delta} = a_t + \Delta s(a_{t+\Delta}). \tag{A.4}$$

Assume that agents dissave, $s(a) \leq 0$ (the case with $s(a) > 0$ is symmetric). Then, from equation (A.4) the fraction of agents with wealth below a at period $t + \Delta$ is:

$$Pr(a_{t+\Delta} \leq a) = Pr(a_t < a_t^o) \cdot Pr(a_t \leq a - \Delta s_t(a)) + Pr(a_t \geq a_t^o) \cdot Pr(a_t \leq a - \Delta s_t(a)),$$
$$\Leftrightarrow \Gamma_{t+\Delta}(a) = \Gamma_t(a_t^o) \cdot \Gamma_t\left(a - \Delta s(a)\right) + (1 - \Gamma_t(a_t^o) \cdot \Gamma_t(a - \Delta s_t(a)).$$

Subtracting $\Gamma_t(a)$ from both sides and dividing by Δ:

$$\frac{\Gamma_{t+\Delta}(a) - \Gamma_t(a)}{\Delta} = \Gamma_t(a_t^o)\left(\frac{\Gamma_t(a - \Delta s_t(a)) - \Gamma_t(a)}{\Delta}\right) + (1 - \Gamma_t(a^o))\left(\frac{\Gamma_t(a - \Delta s_t(a)) - \Gamma_t(a)}{\Delta}\right). \tag{A.5}$$

Note that,

$$\lim_{\Delta \to 0} \frac{\Gamma_t(a - \Delta s_t(a)) - \Gamma_t(a)}{\Delta s_t(a)} \cdot s_t(a) = -d_a\Gamma_t(a) \cdot s_t(a). \tag{A.6}$$

Taking $\Delta \to 0$ in (A.5) and using (A.6) leads to (3.15). ■

A.3 Proof of Lemma 3

Lemma 3

1. *There is a unique stationary tax rate, $\tau^* = r - \rho$.*

2. *The stationary wealth distribution, Γ^*, is non-unique. There is a set of stationary distributions that solves the system:*

$$r - \rho = \frac{b^*\Gamma^*\left(a^{o*}\right)Y(\Gamma^*)}{A^*}, \tag{A.7}$$

$$a^{o*} = \Lambda(\Gamma^*), \tag{A.8}$$

$$b^* = P(\Gamma^*). \tag{A.9}$$

Proof:

Proof of item 1.

Replace $d_t\Gamma_t(a) = 0$ in the KF equation (3.15) and use that $s_t(a) = \theta_t y_t(a)$ to obtain that $\theta^* = 0$. This condition implies that $\tau^* = r - \rho$. Thus, there is a unique tax rate level that is consistent with the economy reaching a steady state wealth distribution.

Proof of item 2.

First, combine equations (3.12) and (3.13) to write the occupational threshold as a function of Γ^*:

$$a^{o*} = \Lambda(\Gamma^*).$$

Evaluate the budget constraint of the government (2.2) at the steady state to obtain:

$$\tau^* = \frac{b^*(1-e^*)y(\Gamma^*)}{A^*}.$$

Use that $e^* = 1 - \Gamma^*(a^{o*})$ and that $\tau^* = r - \rho$ to obtain equation (3.17). Finally, recall that $b^* = P(\Gamma^*)$. Combining the three equations in the lemma gives:

$$r - \rho = \frac{P(\Gamma^*) \cdot \Gamma^*(\Lambda(\Gamma^*)) \cdot y(\Gamma^*)}{A^*}$$

Fix the mean of the distribution A^* to some bounded number in $[\underline{a}, +\infty)$. The condition above restricts the set of distributions consistent with an steady state in the economy. ■

A.4 Proof of Proposition 1

Proposition 1 *Consider an initial wealth distribution Γ_0. Define the initial tax rate $\tau_0 \equiv \tau(\Gamma_0)$ and the initial fraction of entrepreneurs $e_0 \equiv 1 - \Gamma_0(a^o(\Gamma_0))$. Further, restrict Γ_0 to the set of distributions that satisfy $\tau_0 \leq r - \rho$ and $e_0 \geq \frac{\alpha}{2}$. Consider the following three cases for which the transitions dynamics can be completely characterized:*

1. *If $r - \rho < 0$ and $b < 0$, then:*

- *$d_t\tau_t \geq 0$ and $\tau^* = r - \rho$.*

2. *If $r - \rho > 0$ and $b < 0$, then:*

   - *$d_t\tau \geq 0$ and $\tau^* = 0$.*

3. *If $r - \rho > 0$ and $b > 0$, then:*

   - *$d_t\tau \leq 0$ and $\tau^* = 0$.*

All three cases satisfy: i) $d_t\Gamma_t(a) \leq 0\ \forall a$, ii) $d_t a_t^o \geq 0$, iii) $d_t k_t \geq 0$, iv) $\exists \tilde{t} > 0\ :\ k_t = R\tilde{e},\ \forall t \geq \tilde{t}$, and v) $d_t A_t \geq 0$. The steady state fraction of entrepreneurs, $\tilde{e}$, solves $OC(\tilde{e}, b) = 0$ in equation (3.2).

Proof: I start by proving items i) to iv). Then, I show the results regarding the evolution of τ.

Proof of i)

Because I restrict Γ_0 to the set of distributions that satisfy $\tau_0 \leq r - \rho$, then $\theta_t \geq 0, \forall t$. Thus, agents save over time, i.e. $s_t(a) \geq 0, \forall t, \forall a$. The KF equation (3.15) implies that $d_t\Gamma_t(a) \leq 0, \forall t, \forall a$. That is, the wealth distribution shifts right over time in the FOSD sense.

Proof of ii)

First, note that the price of capital, wage rate, and output are given by:

$$p_t = Z\alpha(Re_t)^{\alpha-1}(\ell(1-e_t))^{1-\alpha}, \tag{A.10}$$

$$w_t = Z(1-\alpha)(Re_t)^{\alpha}(\ell(1-e_t))^{-\alpha}, \tag{A.11}$$

$$Y_t = Z(Re_t)^{\alpha}(\ell(1-e_t))^{1-\alpha}. \tag{A.12}$$

Differentiation of (A.10) to (A.12) gives:

$$dp_t = -\alpha(1-\alpha)Z\left[R(Re_t)^{\alpha-2}(\ell(1-e_t))^{1-\alpha} + \ell(Re_t)^{\alpha-1}(\ell(1-e_t))^{-\alpha}\right]de_t$$
$$= -(1-\alpha)p_t\left[\frac{1}{e_t} + \frac{1}{1-e_t}\right]de_t \tag{A.13}$$
$$dw_t = \alpha(1-\alpha)Z\left[R(Re_t)^{\alpha-1}(\ell(1-e_t))^{-\alpha} + \ell(Re_t)^{\alpha}(\ell(1-e_t))^{-\alpha-1}\right]de_t$$
$$= \alpha w_t\left[\frac{1}{e_t} + \frac{1}{1-e_t}\right]de_t \tag{A.14}$$
$$dY_t = Z\left[\alpha R(Re_t)^{\alpha-1}(\ell(1-e_t))^{1-\alpha} + (1-\alpha)\ell(Re_t)^{\alpha}(\ell(1-e_t))^{-\alpha}\right]de_t$$
$$= [p_tR - w_t\ell]\,de_t \tag{A.15}$$

Note that the sign of dY_t depends on $p_tR - w_t\ell$. At any given period t, the occupational constraint implies that $p_tR - w_t\ell \geq rI + bY_t$. Assumption 1 on the minimum transfer rate -$\underline{b}$ guarantees that $rI + bY_t \geq 0$, and thus, $p_tR - w_t\ell \geq 0$. Hence, sign$\{dY_t\}$ = sign$\{de_t\}$. Additionally, sign$\{dp_t\}$ = sign$\{-de_t\}$ and sign$\{dw_t\}$ = sign$\{de_t\}$.

In what follows I use (A.13) to (A.15) to find the evolution of the occupational threshold over time, i.e. $d_ta_t^o$. I start by considering the case in which the IC binds, and then I study the case when the OC binds.

Case 1: IC binds

First, suppose that the IC constraint binds ($IC_t = 0 \Rightarrow dIC_t = 0$). Then,

$$(1+r)d\hat{a}_t = dw_t\ell + dbY + bdY - dpR,$$
$$= \left[w\ell\left(\alpha\left(\frac{1}{e_t} + \frac{1}{1-e_t}\right) - b\right) + pR\left((1-\alpha)\left(\frac{1}{e_t} + \frac{1}{1-e_t}\right) + b\right)\right]d_te_t + dbY_t. \tag{A.16}$$

The term in the square brackets is positive by Assumption 1. Call this term χ_t and note that $d_te_t = -d_t(\Gamma_t(a_t^o))$. Because the IC binds, then $a_t^o = \hat{a}_t$. Therefore, equation (A.16) reads as:

$$(1+r)d_t\hat{a}_t = -\chi_t d_t(\Gamma_t(\hat{a}_t)) + d_tbY_t. \tag{A.17}$$

Imposing that $b_t = b$, noting that $d_t(\Gamma_t(\hat{a}_t)) = d_t\Gamma_t(\hat{a}_t) + d_a\Gamma_t(\hat{a}_t)d_t\hat{a}_t$, and evaluating the derivative in terms of t gives:

$$\begin{aligned}(1+r)d_t\hat{a}_t &= -\chi_t\left(d_t\Gamma_t(\hat{a}_t) + d_a\Gamma_t(\hat{a}_t)d_t\hat{a}_t\right),\\ \Rightarrow d_t\hat{a}_t &= -\frac{\chi_t d_t\Gamma_t(\hat{a}_t)}{1+r+\chi_t d_a\Gamma_t(\hat{a}_t)}.\end{aligned} \tag{A.18}$$

Because $d_a\Gamma_t(\hat{a}_t) \geq 0$, equation (A.18) implies that: $\text{sign}\{d_t\hat{a}_t\} = \text{sign}\{-d_t\Gamma_t(\hat{a}_t)\}$. Further, the KF equation (3.15) implies that $\text{sign}\{-d_t\Gamma_t(\hat{a}_t)\} = \text{sign}\{\theta_t\}$, thus $\text{sign}\{d_t\hat{a}_t\} = \text{sign}\{\theta_t\}$. Because $\theta_t \geq 0$, then $d_t\hat{a}_t \geq 0$.

Case 2: OC binds

Now suppose that the OC binds ($OC_t = 0 \Rightarrow d_tOC_t = 0$). Proceeding as in Case 1 leads to:

$$\left[w\ell\left(\alpha\left(\frac{1}{e_t}+\frac{1}{1-e_t}\right)-b\right)+pR\left((1-\alpha)\left(\frac{1}{e_t}+\frac{1}{1-e_t}\right)+b\right)\right]d_te_t + d_tbY_t = 0. \tag{A.19}$$

In this case, the occupational threshold is $a_t^o = \tilde{a}_t$. Thus, equation (A.19) reads as:

$$\chi_t d_t(\Gamma_t(\tilde{a}_t)) + d_tbY = 0 \tag{A.20}$$

Using that $b_t = b$ and evaluating the derivative at t gives:

$$\begin{aligned}\chi_t\left(d_t\Gamma_t(\tilde{a}_t) + d_a\Gamma_t(\tilde{a}_t)d_t\tilde{a}_t\right) &= 0,\\ \Rightarrow d_t\tilde{a}_t &= -\frac{d_t\Gamma_t(\tilde{a}_t)}{d_a\Gamma_t(\tilde{a}_t)}.\end{aligned} \tag{A.21}$$

Analogously to Case 1, equation (A.21) implies that $\text{sign}\{\tilde{a}_t\} = \text{sign}\{\theta_t\}$. Since $\theta_t \geq 0$, then $d_t\tilde{a}_t \geq 0$.

Recall that the OC constraint starts to bind when $\hat{a}_t \geq \frac{I}{1+r}$. Thus, combining both cases, the evolution of the occupational threshold a_t^o is given by:[35]

$$d_t a_t^o = \begin{cases} -\frac{\chi_t d_t \Gamma_t(\hat{a}_t)}{1+r+\chi_t d_a \Gamma_t(\hat{a}_t)} & \text{if } \hat{a}_t \leq \frac{I}{1+r}, \\ -\frac{d_t \Gamma_t(\tilde{a}_t)}{d_a \Gamma_t(\tilde{a}_t)} & \text{if } \hat{a}_t > \frac{I}{1+r}. \end{cases} \tag{A.22}$$

Proof of item iii)

Equation (A.17) and (A.20) evaluated at t imply that:

$$d_t e_t = \begin{cases} \frac{(1+r)}{\chi_t} d_t \hat{a}_t & \text{if } \hat{a}_t \leq \frac{I}{1+r}, \\ 0 & \text{if } \hat{a}_t > \frac{I}{1+r}. \end{cases} \tag{A.23}$$

From item ii), I conclude that $d_t e_t \geq 0$. Thus, $d_t k_t = R d_t e_t \geq 0$.

Proof of item iv)

From item ii), the effective occupational threshold increases over time, thus it reaches $\frac{I}{1+r}$ in a finite number of periods $\tilde{t}$. Therefore, the occupational constraint binds for $t \geq \tilde{t}$ and so $e_t = \tilde{e}, \forall t \geq \tilde{t}$. As a result, $k_t = R\tilde{e}$ for $t \geq \tilde{t}$.

Proof of item v)

Because $\theta_t \geq 0$, equation (3.9) implies that $d_t A_t \geq 0$.

Now I proceed to show the results regarding the evolution of τ. Differentiate the government budget constraint (2.2) in terms of t:

$$\begin{aligned} d_t \tau_t A_t + \tau_t d_t A_t &= b(d_t Y_t (1-e_t) - Y_t d_t e_t), \\ &= b(\,(pR - w\ell)(1-e) - Y_t\,) d_t e_t, \\ &= bY_t \left(\frac{\alpha(1-e_t)}{e_t} - (1-\alpha) - 1 \right) d_t e_t, \\ &= bY_t \frac{\alpha - 2e_t}{e_t} d_t e_t, \end{aligned}$$

[35]Note that equation (A.22) predicts that there is a kink in $d_t a_t^o$ at $\hat{a} = \frac{I}{1+r}$. This feature is illustrated by the third row of graphs in Figure 6.

where in the second line I have used (A.15). From the expression above, the evolution of τ is given by

$$d_t\tau_t = \frac{1}{A_t}\left(bY_t\frac{[\alpha - 2e_t]}{e_t}\right)d_te_t - \tau_t d_t A_t. \tag{A.24}$$

Note first that in the proposition I restrict Γ_0 such that $e_0 \geq \frac{\alpha}{2}$. From item iii), $d_te_t \geq 0$, thus the term in brackets in equation (A.24) is always negative.

Consider now Case 1. Because $b < 0$, then also $\tau_t < 0$. As long as $\theta_t > 0$, $d_tA_t > 0$ and so $d_t\tau_t > 0$. The tax rate increases until it reaches its steady state level $\tau^* = r - \rho$ which makes the saving rate equal to zero, and thus, $\lim_{t\to+\infty} d_tA_t = 0$.

Consider Case 2. As in Case 1, $b < 0$ so τ_t increases over time. However, because $b < 0$, the tax rate is bounded from above by zero. Thus, it will never attain the rate that is required to keep the wealth distribution constant: $r - \rho > 0$. Hence, $\theta_t > 0, \forall t$ and thus, $d_tA_t > 0, \forall t$. Because wealth keeps increasing over time, the budget constraint of the government (2.2) implies that $\lim_{t\to+\infty} \tau_t = 0$, i.e. $\tau^* = 0$.

Finally, consider Case 3. In this case, $b > 0$ so $\tau_t \geq 0$. Equation (A.24) implies now that $d_t\tau_t \leq 0$. Because $\tau_0 \leq r - \rho$ and the tax rate decreases over time, it will never attain $r - \rho$. As in Case 2, the economy accumulates wealth over time and so the budget constraint of the government implies that $\tau^* = 0$. ■

A.5 Proof of Lemma 4

Lemma 4 *Given some wealth distribution Γ_t, a marginal increase in the transfer rate, b, leads to: i) an increase of the minimum collateral $\hat{a}$, ii) an increase of the occupational threshold $\tilde{a}$, iii) a decrease of the wage rate w, and iv) and increase of the price of capital p.*

Proof: Equations (A.17) and (A.20) imply that:

$$d_b\hat{a} = \frac{Y}{1 + r + \chi\gamma(\hat{a})} > 0,$$
$$d_b\tilde{a} = \frac{Y}{\chi\gamma(\tilde{a})} > 0.$$

Therefore, $d_b a^o > 0$, which implies that $d_b e =< -\gamma(a^o) d_b a^o$. Thus, equations (A.13) and (A.14) imply that: $d_b p > 0$ and $d_b w < 0$. ■

A.6 Proof of Lemma 5

Lemma 5 *The preferred transfer rate function* $b(a, \Gamma)$ *is as follows:*

$$b(a; \Gamma) = \begin{cases} \psi^1(a; \Gamma) & \text{if } a < \hat{a}(\text{-}\underline{b}, \Gamma), \\ \psi^2(a; \Gamma) & \text{if } a \in [\hat{a}(\text{-}\underline{b}, \Gamma), \overline{a}(\Gamma)], \\ \psi^3(a; \Gamma) & \text{if } a \geq \overline{a}(\Gamma), \end{cases}$$

where the functions $\psi^j(a; \Gamma)$, $j \in \{1, 2, 3\}$, *are continuous in assets and satisfy:*

$d_a \psi^1 \leq 0$, $d_a \psi^2 > 0$, $d_a \psi^3 \leq 0$, $\psi^2(\hat{a}(\text{-}\underline{b}; \Gamma)) < \lim_{a \to \hat{a}(\text{-}\underline{b}, \Gamma)^-} \psi^1(a; \Gamma)$ *and* $\lim_{a \to +\infty} \psi^3(a; \Gamma) = \text{-}\underline{b}$. *Further,* $\overline{a}(\Gamma) \in (\hat{a}(\text{-}\underline{b}, \Gamma), \tilde{a}(\overline{b}, \Gamma)]$.

Proof:

In what follows, I omit the dependence on Γ to simplify notation. Figure 14 in section C.4 illustrates the disposable income function that can be useful to follow this proof. I organize the proof according to the preferences of the three identified endogenous classes.

Working Class

First, because the minimum collateral is increasing in b, $\hat{a}(\text{-}\underline{b})$ is the minimum possible requirement to start a firm. Therefore, any agent with $a < \hat{a}(\text{-}\underline{b})$ cannot start a firm and remains a worker (working class, WC). Thus, agents with $a < \hat{a}(\text{-}\underline{b})$ solve:

$$b(a) = \arg\max_{b \in [-\underline{b}, \overline{b}]} \{\, y^W(b, a) \equiv (r - \tau(b))a + w(b)\ell + T(b) \,\}. \tag{A.25}$$

Differentiating the income of workers in terms of b gives:

$$d_b y^W = d_b w \cdot \ell + d_b T - d_b \tau \cdot a.$$

Further, the budget condition of the government implies:

$$d_b T = \frac{1}{1-e}(A\, d_b\tau + T d_b e).$$

Combining both conditions,

$$d_b y^W = d_b w \cdot \ell + d_b\tau \left(\frac{A}{1-e} - a\right) + \frac{T}{1-e} d_b e. \tag{A.26}$$

A similar procedure that was used to obtain equation (A.24) can be used to show that

$$d_b\tau = y\left((1-e) + b\left(\frac{\alpha - 2e}{e}\right) d_b e\right)\frac{1}{A} \geq 0,$$

Differentiation of (A.26) in terms of a gives: $d_a(d_b y^W) = -d_b\tau \leq 0$. Thus, $y^W(a,b)$ is submodular in (a,b). If problem (A.25) has a unique solution, then the Topkis' univariate theorem implies that $d_a b(a) \leq 0$ for $a < \hat{a}(\text{-}\underline{b})$. That is, the preferred transfer rate is characterized by a weakly decreasing function in a, denoted by $\psi^1(a)$ in the lemma.[36]

Emerging Class

Consider now agents with $a \in [\hat{a}(\text{-}\underline{b}), \tilde{a}(\overline{b}))$. The OC does not bind in this range. Therefore, $\Pi > w\ell + T$, which implies that $y^E(a,b) > y^W(a,b)$, where y^E is the entrepreneur's income which is given by: $y^E(a,b) \equiv (r - \tau(b))a + \Pi(b)$. Agents with $a \geq \hat{a}(\text{-}\underline{b})$ decide to become entrepreneurs, and thus, solve:

$$\max_{b \in [-\underline{b}, \overline{b}]} \{\, y^E(a,b) \,\}$$
$$s.t.\ a \geq \hat{a}(b). \tag{A.27}$$

Consider the unconstrained solution to problem (A.27), denoted by $b^E(a)$. Because $d_a(d_b y^E)) = -d_b\tau \leq 0$, the Topkis' theorem implies that $d_a b^E(a) \leq 0$. Consider the maximum transfer rate

[36] If problem (A.25) does not have a unique solution, then if $a' > a$ the solution satisfies $\psi^1(a') \leq \psi^{WC}(a)$ (in the strong set order).

that allows an agent with assets a to start a firm: $\tilde{b}(a)\ :\ a = \hat{a}(\tilde{b})$. This transfer rate satisfies $d_a\tilde{b} = \frac{1}{d_b\hat{a}} > 0$. Figure 15 depicts $\tilde{b}(a)$.

Define the auxiliary function $\Theta(a) \equiv \tilde{b}(a) - b^E(a)$, which is continuous and increasing in assets. First, note that: $\Theta(\ \hat{a}(\text{-}\underline{b})\ ) = \text{-}\underline{b} - b^E(\ \hat{a}(\text{-}\underline{b})\ ) \leq 0$. Further, because Θ is continuous and increasing, there exists a unique $\overline{a} \in [\hat{a}(\text{-}\underline{b}), \hat{a}(\overline{a})]$ such that $\Theta(\overline{a}) = 0$. Agents with assets $a \in (\hat{a}, \overline{a}]$ are limited by the incentive compatibility constraint. Thus, their preferred transfer rate is given by $\tilde{b}(a)$ which is increasing in a. This function is denoted by $\psi^2(a)$ in the lemma (emerging class).

Incumbent Class

Finally, if $a > \overline{a}$, then the solution to problem (A.27) is given $b^E(a)$ which is weakly decreasing and denoted by the function $\psi^3(a)$ in the lemma (incumbent class). Additionally, $d_b y^E(a) = \underbrace{d_b p}_{>0} R - \underbrace{d_b \tau}_{\geq 0} a$. Thus, $\lim_{a\to+\infty} d_b y^E(a) < 0$ which implies that $\lim_{a\to+\infty} \psi^3(a) = \text{-}\underline{b}$. ■

A.7 Proof of Lemma 6

Lemma 6 *There is a unique transfer rate, b_t, that solves the government's problem* (5.3)*:*

$$1 - \Gamma_t(a^o(b_t, \Gamma_t)) = e^*,$$

where $e^* < \alpha$ is a function of aggregate productivity and the fixed parameters of the model.

Proof: Rewrite the welfare function of the government as follows:

$$\mathcal{W}(b, \Gamma_t) = (1 - \alpha(1 - \phi))Y_t - \phi\frac{rI}{R}K_t.$$

Taking the first order condition gives:[37]

$$\left[(1-\alpha(1-\phi))\left(\alpha RK_t^{\alpha-1}L_t^{1-\alpha}-(1-\alpha)\ell K_t^{\alpha}L_t^{-\alpha}\right)-\phi rI\right]\cdot\underbrace{d_b e_t}_{<0}=0,$$

$$\Leftrightarrow (pR-w\ell)=\frac{\phi}{(1-\alpha(1-\phi))}rI. \tag{A.28}$$

Alternatively, the previous condition can be rewritten as

$$\alpha e_t^{\alpha-1}(1-e_t)^{1-\alpha}-(1-\alpha)e_t^{\alpha}(1-e_t)^{-\alpha}=\frac{\phi}{(1-\alpha(1-\phi))}\frac{rI}{ZR^{\alpha}\ell^{1-\alpha}}. \tag{A.29}$$

Call the function on the left-hand side $\tilde{h}(e)$ and the expression on the right-hand side $\tilde{\phi}$. Note that $\tilde{h}(e)$ is strictly decreasing in e. Additionally, $\lim_{e\to 0}\tilde{h}(e)=+\infty$ and $\lim_{e\to 1}\tilde{h}(e)=-\infty$. Since $\tilde{h}(e)$ is continuous, by the intermediate value theorem there is an optimal fraction of entrepreneurs, e^* such that $\tilde{h}(e^*)=\tilde{\phi}$. Moreover, because $\tilde{h}$ is strictly decreasing, e^* is unique.

Additionally, if $e=\alpha$, then $\tilde{h}=0$. Because $\tilde{h}$ is decreasing in e and $\tilde{\phi}>0$ it must be that $e^*<\alpha$. Figure 26 illustrates the features of the political equilibrium condition (A.29). This concludes the proof. ■

A.8 Proof of Corollary 1

Corollary 1 *Define* $\varphi=\frac{(\phi-1)(1-\alpha)}{1-\alpha(1-\phi)}$ *and the political equilibrium output:* $Y_{PE}=Z(Re^*)^{\alpha}(\ell(1-e^*))^{1-\alpha}$.

1. *The effective occupational threshold is given by:* $a_t^o=\Gamma_t^{-1}(1-e^*)$.

2. *The occupational and incentive compatibility functions read as:*

$$OC(b_t)=\varphi rI-b_tY_{PE},$$

$$IC(b_t)=\varphi rI-b_tY_{PE}-[I-(1+r)\Gamma_t^{-1}(1-e^*)].$$

[37]Note that from (A.28): $\Pi_t=pR-rI=\frac{(\phi-1)(1-\alpha)}{1-\alpha(1-\phi)}rI+w\ell$. Thus, having that $\phi>1$ is a sufficient condition to have that $\Pi_t>w_t\ell$, so that the occupational constraint can hold at the desired transfer rate b_t. Otherwise, the government would not be able to attain the optimal fraction of entrepreneurs.

3. *The equilibrium transfer rate b_t is given by:*

$$b_t = \begin{cases} \frac{(\varphi r-1)I+(1+r)a_t^o}{Y_{PE}} & \text{if } a_t^o \leq \frac{I}{1+r}, \\ \overline{b} \equiv \frac{\varphi r I}{Y_{PE}} & \text{if } a_t^o > \frac{I}{1+r}. \end{cases}$$

Proof:

Proof of item 1.

Inverting the PE condition (5.4) gives the result.

Proof of item 2.

Equation (A.28) implies that $pR - rI - w\ell = \frac{(\phi-1)(1-\alpha)}{1-\alpha(1-\phi)} rI = \varphi r I$. Replacing in equation (3.2) and using that $Y_{PE} = Z(Re^*)^\alpha(\ell(1-e^*))^{1-\alpha}$ gives equation (5.5). To obtain equation (5.6) use the previous result and that $a_t^o = \Gamma_t^{-1}(1-e^*)$.

Proof of item 3.

The OC and IC intersect at $a_t^o = \frac{I}{1+r}$. Thus, when $a_t^o \leq \frac{I}{1+r}$, the transfer rate is obtained by solving $IC(b_t) = 0$. Otherwise, when $a_t^o > \frac{I}{1+r}$, the OC binds and so $b_t = \overline{b}$. This leads to equation (5.7). ∎

A.9 Proof of Proposition 2

Proposition 2 *Consider an initial wealth distribution, Γ_0. Define the initial transfer rate $b_0 = P(\Gamma_0)$ and the initial tax rate $\tau_0 = \tau(\Gamma_0)$. Further, denote the long-run value of some variable x by $x^* = \lim_{t\to+\infty} x_t$. Then, the dynamics of the political equilibrium are as follows:*

1. *If $\tau_0 < r - \rho$, then: $d_t b_t \geq 0$, $d_t A_t \geq 0$, and $d_t a_t^o \geq 0$.*

   (a) *If $r - \rho < 0$, then: $b^* \in (-\underline{b}, 0)$, $\tau^* = r - \rho$, and $A^* \in (\underline{a}, +\infty)$.*

   (b) *If $r - \rho > 0$, there are two cases:*

      i. *$b^* \in (0, \overline{b}]$, $\tau^* = r - \rho$, and $A^* \in (\underline{a}, +\infty)$.*

      ii. *$b^* = \overline{b}$, $\tau^* = 0$, and $A^* \to +\infty$.*

2. *If* $\tau_0 > r-\rho$, *then:* $d_t b_t \leq 0$, $d_t A_t \leq 0$, *and* $d_t a_t^o \leq 0$.

   (a) *If* $r-\rho<0$ *and* $b_0<0$, *then:* $b^* \in [\text{-}\underline{b}, 0)$, $\tau^* = r-\rho$, *and* $A^* \in (\underline{a}, +\infty)$.

   (b) *If* $b_0 > 0$, *there are two cases:*

      i. $b^* \in [\text{-}\underline{b}, \overline{b}]$, $\tau^* = r-\rho$, *and* $A^* \in (\underline{a}, +\infty)$.

      ii. $b^* = \text{-}\underline{b}$, $\tau^* = \frac{-\underline{b}\, y(e^*)(1-e^{*\prime})}{\underline{a}}$, *and* $A^* = \underline{a}$, *where* $e^{*\prime}$ *solves* $h(e^{*\prime}) = (1+r)(I-\underline{a})$.

Proof: I start by showing items 1. and 2. Then, I show items (a) and (b) from each case. Differentiate the PE condition (A.29) to obtain:

$$\begin{aligned}
-\alpha(1-\alpha)\underbrace{\left(e_t^{\alpha-2}(1-e_t)^{1-\alpha} + 2e_t^{\alpha-1}(1-e_t)^{-\alpha} + e_t^{\alpha}(1-e_t)^{-1-\alpha}\right)}_{>0} d_t e_t &= 0 \\
\Rightarrow d_t e_t = \frac{d}{dt}\left(1-\Gamma_t(a_t^o)\right) &= 0 \\
d_t\Gamma_t(a_t^o) + d_a\Gamma_t(a_t^o) d_t a_t^o &= 0 \qquad \text{(A.30)}
\end{aligned}$$

Equations (A.17) and (A.20) imply that:

$$d_t\hat{a}_t = \frac{Y_t d_t b_t - \chi_t d_t\Gamma_t(\hat{a}_t)}{1+r+\chi_t d_a\Gamma_t(\hat{a}_t)}, \qquad \text{(A.31)}$$

$$d_t\tilde{a}_t = \frac{Y_t d_t b_t - \chi_t d_t\Gamma_t(\tilde{a}_t)}{\chi_t d_a\Gamma_t(\tilde{a}_t)}. \qquad \text{(A.32)}$$

The OC starts to bind when $\hat{a}_t \geq \frac{I}{1+r}$. Thus, combining both equations gives:

$$d_t a_t^o = \begin{cases} \frac{Y_t d_t b_t - \chi_t d_t\Gamma_t(\hat{a}_t)}{1+r+\chi_t d_a\Gamma_t(\hat{a}_t)} & \text{if } \hat{a}_t \leq \frac{I}{1+r} \\ \frac{Y_t d_t b_t - \chi_t d_t\Gamma_t(\tilde{a}_t)}{\chi_t d_a\Gamma_t(\hat{a}_t)} & \text{if } \hat{a}_t > \frac{I}{1+r} \end{cases} \qquad \text{(A.33)}$$

Suppose first that $\hat{a}_t \leq \frac{I}{1+r}$, then combining equations (A.30) and (A.33) leads to

$$d_t\Gamma_t(\hat{a}_t)\left(1 - \frac{d_a\Gamma_t(\hat{a}_t)\chi_t d_t\Gamma_t(\hat{a}_t)}{1+r+\chi_t d_a\Gamma_t(\hat{a}_t)}\right) = \frac{d_a\Gamma_t(\hat{a}_t)Y_t d_t b_t}{1+r+\chi_t d_a\Gamma_t(\hat{a}_t)},$$
$$\Rightarrow d_t b_t = \frac{-d_t\Gamma_t(\hat{a}_t)}{d_a\Gamma_t(\hat{a}_t)}\frac{1+r}{Y_t}.$$

Similarly, if $\hat{a}_t > \frac{I}{1+r}$, then $d_t b_t = 0$. Combining both cases gives:

$$d_t b_t = \begin{cases} \frac{-d_t\Gamma_t(\hat{a}_t)}{d_a\Gamma_t(\hat{a}_t)}\frac{1+r}{Y_t} & \text{if } \hat{a}_t \leq \frac{I}{1+r}, \\ 0 & \text{if } \hat{a}_t > \frac{I}{1+r}. \end{cases} \tag{A.34}$$

Therefore, from the KF equation (3.15), $\text{sign}\{b_t\} = \text{sign}\{-d_t\Gamma_t\} = \text{sign}\{\theta_t\}$. Moreover, equation (A.33) implies that $\text{sign}\{d_t a_t^o\} = \text{sign}\{\theta_t\}$. Hence, if $\tau_0 < r - \rho$, then $\theta_0 > 0$ which implies that $db_t \geq 0$, $d_t a_t^o \geq 0$, and $d_t A_t \geq 0$. Otherwise, these dynamics are reversed.

Now I proceed to show items a) and b) of each case. Use the PE condition (5.4) to write the budget constraint of the government as follows

$$\tau_t A_t = b_t \underbrace{Y_{PE}(1-e^*)}_{\equiv\Omega}.$$

Differentiate in terms of t to obtain

$$d_t\tau_t = \frac{1}{A_t}\left(d_t b_t \Omega - \tau_t d_t A_t\right) \tag{A.35}$$

Case 1. (a): because $r - \rho < 0$ and $\tau_0 < r - \rho$, then $\tau_0 < 0$. Equation (A.35) implies that $d_t\tau_t \geq 0, \forall t$. Thus, $\tau^* = r - \rho < 0$, which implies that $b^* \in (-\underline{b}, 0)$. Also, because $\theta^* = 0$, the economy reaches a stationary distribution with some mean $A^* \in (\underline{a}, +\infty)$.

Case 1. (b): because $r - \rho > 0$, the tax rate can be either negative or positive along the transition path. If $\tau_t \leq 0$, then equation (A.35) gives $d_t\tau_t > 0$. However, when $\tau_t \geq 0$, the sign of $d_t\tau_t$ is ambiguous. As stated in the proposition, there are two cases: i. The tax rate reaches $\tau^* = r - \rho > 0$. Thus, it must be that $b^* \in (0, \bar{b}]$. ii. The transfer rate reaches its maximum before the tax rate can attain $r - \rho$. Thus, $b^* = \bar{b}$ and $\theta^* > 0$ which implies that $A^* \to +\infty$. The budget constraint of the government implies that $\tau^* \to 0$.

Case 2. (a): because $r - \rho < 0$ and $b_0 < 0$, then $\tau_0 < 0$. Equation (A.35) implies that $d_t\tau_t \leq 0, \forall t$. Thus, the proof proceeds as in Case 1. (a).

Case 2. (b): because $b_0 > 0$, then $\tau_0 > 0$. Regardless of the sign of $r-\rho$, the tax rate can be either positive or negative along the transition path which implies that the sign of $d_t\tau_t$ can be ambiguous. Thus, as in Case 2. (b). there are two cases. i. The tax rate reaches $\tau^* = r - \rho$. Depending on the sign of $r - \rho$ the stationary transfer rate can be positive or negative, thus $b^* \in [\text{-}\underline{b}, \bar{b}]$. Also, because $\theta^* = 0$, the economy reaches a stationary distribution with mean $A^* > \underline{a}$.

ii. The transfer rate reaches its minimum before the tax rate can attain $r - \rho$. Thus, $b^* = \text{-}\underline{b}$ and $\theta^* < 0$, which implies that $A^* \to \underline{a}$. The behavior of τ is more involved. First, note that when b_t reaches $\text{-}\underline{b}$, the IC constraint remains binding and so the minimum collateral to get a loan $\hat{a}_t$ defines the fraction of entrepreneurs given Γ_t. As a result, from that point onwards, e_t is different from the optimal fraction of entrepreneurs e^*. Because $d_t\hat{a}_t \leq 0$, equation (A.23) implies that $d_te_t \leq 0$. The evolution of τ_t is given by equation (A.24):

$$d_t\tau_t = \frac{1}{A_t}\left(Y_t\frac{[\alpha - 2e_t]}{e_t}\right)\underbrace{\text{-}\underline{b}\,d_te_t}_{>0}\underbrace{-\tau_t d_tA_t}_{>0}.$$

Note that when $e_t > \frac{\alpha}{2}$, the sign of $d_t\tau_t$ is ambiguous. Thus, the tax rate may keep decreasing for a while after the economy reaches $\text{-}\underline{b}$. However, because $d_te_t \leq 0$, at some point $e_t < \frac{\alpha}{2}$ and so there is a t such that τ_t starts increasing over time. In the steady state it must be that $\hat{a}_t = \underline{a}$ which implies that $\tau^* = \frac{-\underline{b}\,y(e^*)(1-e^*)}{\underline{a}}$, where e^* solves $h(e^*) = (1 + r)(I - \underline{a})$. ■

A.10 Proof of Proposition 3

Proposition 3 *Consider some stationary wealth distribution* $\Gamma^* \in \mathcal{G}$ *with mean* A*. The initial wealth distribution is such that* $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$*, with occupational threshold* $a_0^o = a^o(\Gamma_0)$*. Denote by* $\tilde{A} \in N(A)$ *the intersect of both cumulative distributions. Then,*

1. *If* $\tilde{A} < a_0^o$*: i)* $d_t b_t \leq 0$*, ii)* $d_t A_t \leq 0$*, iii)* $d_t a_t^o \leq 0$*.*

2. *If* $\tilde{A} > a_0^o$*, the signs of i) to iii) are reversed.*

The transition dynamics are reversed if $\Gamma_0 \in \mathcal{G}_{MPS_{rev}}(\Gamma^*)$*.*

Proof: Consider an economy that starts with the stationary wealth distribution $G^* \in \mathcal{G}$. There is an unanticipated shock that shifts the wealth distribution according to an MPS as in Definition 1. Take the MPS distribution $\Gamma \in \mathcal{G}_{MPS}(\Gamma^*)$ and write the new distribution Γ_0 after the MPS shock as follows:

$$\Gamma_0 = \theta\Gamma + (1-\theta)\Gamma^*,\ \theta \in [0,1]. \tag{A.36}$$

Marginal increases in θ starting from $\theta = 0$ generate a sequence of MPS distributions of Γ^*, that is $\Gamma_0 \in \mathcal{G}_{MPS}(\Gamma^*)$. In general, when the economy reaches a stationary distribution, the effective occupational threshold satisfies: $a^o = \hat{a}$.[38] Differentiation of $\hat{a}$ in terms of θ gives

$$d_\theta \hat{a} = \frac{d_\theta bY - \chi d_\theta(\Gamma_0(\hat{a}))}{1+r},$$

where I have used equation (A.17). Also, differentiating the first order condition of the government (5.4) gives

$$d_\theta \Gamma_0(\hat{a}) + d_a \Gamma_0(\hat{a}) d_\theta \hat{a} = 0$$

[38] Recall that when the OC constraint binds ($a^o = \tilde{a}$) the economy continues to accumulate wealth over time and does not reach a stationary distribution. The only exception is when the economy reaches $\tau^* = r - \rho$ at the exact same time at which the OC binds, which is a very unlikely event. Thus, in general, if the economy reaches a steady state distribution the IC binds, $a^o = \hat{a}$.

where I have used that $d_\theta(\Gamma_0(\hat{a})) = d_\theta\Gamma_0(\hat{a}) + d_a\Gamma_0(\hat{a})d_\theta\hat{a}$. Combining both conditions leads to

$$\begin{aligned} d_\theta b &= \frac{-d_\theta\Gamma_0(\hat{a})(1+r)}{d_a\Gamma_0(\hat{a})Y}, \\ &= \frac{(\Gamma^*(\hat{a}) - \Gamma(\hat{a}))(1+r)}{d_a\Gamma_0(\hat{a})Y}. \end{aligned} \tag{A.37}$$

Therefore, $\text{sign}\{d_\theta b\} = \text{sign}\{\Gamma^*(\hat{a}) - \Gamma(\hat{a})\}$. From Definition 1, if $\tilde{A} < \hat{a}$ then $d_\theta b > 0$. Otherwise, $d_\theta b < 0$. The budget constraint of the government (3.14) implies that $\tau > r - \rho$ if $\tilde{A} < \hat{a}$. Thus, Proposition 2 implies that: i) $d_t b_t \leq 0$, ii) $d_t A_t \leq 0$, iii) $d_t a_t^o \leq 0$. These dynamics are reversed if $\tilde{A} > \hat{a}$. Finally, all the results are also reversed if $\Gamma_0 \in \mathcal{G}_{MPS_{rev}}(\Gamma^*)$. ■

B Appendix: Extensions

B.1 Political mechanism

In this section, I present a politico-economy mechanism that rationalizes problem 5.3 that defines the equilibrium size of the welfare state, presented in Section 5. I embed the basic environment into a political economy framework with probabilistic voting along the lines of Persson and Tabellini (2000).

Several papers implement probabilistic voting in dynamic models to study the political support for different types of policies. Hassler et al. (2005) study income redistribution, while Gonzalez-Eiras and Niepelt (2008) and Sleet and Yeltekin (2008) focus on social security. Farhi et al. (2012) analyze taxation and Song et al. (2012) examine fiscal policy.

B.1.1 Politicians

Each period t, the electoral competition takes place between two office-seeking candidate governments, A and B. Both candidate governments simultaneously and noncooperatively announce their electoral platforms, $b_t^A \in [\underline{b}, \bar{b}]$ and $b_t^B \in [\underline{b}, \bar{b}]$. At the beginning of each period, candidates observe the wealth distribution Γ_t and make their announcements to maximize their share of votes. candidates' only relevant state variable is the current wealth distribution, Γ_t. There are new elections every period, thus candidates cannot make credible commitments over future social benefits. Probabilistic voting smooths the political objective function by introducing uncertainty from the candidates' point of view (Lindbeck and Weibull, 1987). The specific sources of uncertainty are described in next section.

B.1.2 Voters

At the beginning of each period t, agents observe their assets a and the wealth distribution, Γ_t. Given the state (a, Γ_t), agents can anticipate the effects that a given transfer rate has on occupational choice. They vote before making consumption-saving decisions and to maximize a weighted measure of income:[39]

$$\bar{y}_t(a,b) = ra + \begin{cases} w_t(b)\ell + \frac{T_t(b)-\tau_t(b)a}{\phi^W} & \text{if } a < a_t^0(b), \\ \Pi_t(b) - \frac{\tau_t(b)a}{\phi^E} & \text{if } a \geq a_t^0(b), \end{cases} \tag{B.1}$$

where $\phi^j > 0,\ j \in \{W,E\}$ are occupation-specific preference parameters. When $\phi^j = 1$, then $\bar{y}$ is equal to disposable income. The individual preferences generated by $\bar{y}$ are qualitatively similar to those studied in Section 5.2. The normalization used in this section simplifies the aggregation of preferences. The parameter ϕ^W measures the workers' "taste" for social benefits, while ϕ^E measures how much entrepreneurs dislike paying taxes. I assume that $\phi^E > \phi^W$, thus entrepreneurs dislike more paying taxes than what workers like receiving benefits.[40]

Voters have heterogeneous political preferences not only over redistribution as measured by b, but also over other policy dimensions that are orthogonal to b. Specifically, each period t each voter with assets a has ideological preferences denoted by v_t. Thus, each period there is a continuum of agents indexed by (a, v_t).

Voter (a, v_t) votes for candidate A if:

$$\bar{y}_t(a, b_t^A) > \bar{y}_t(a, b_t^B) + v_t, \tag{B.2}$$

where v_t corresponds to the ideological preference for candidate B. That is, the inherent bias of a voter in period t for party B, irrespective of the proposed policy platforms. The distribution of v_t is assumed to be uniform on $[-1/(2\phi^j), 1/(2\phi^j)],\ j \in \{W,E\}$. Candidates know the ideological distributions before announcing their policy platforms. However, they do not know the specific realizations of ideological preferences. Therefore, candidates announce their policies under uncertainty about the results of the election.

[39] In what follows, the time subscript captures the dependence of the endogenous variables on the wealth distribution Γ_t.

[40] This assumption guarantees that in equilibrium $\Pi_t > w_t\ell$, so the occupational constraint can hold for a set of possible transfer rates. Otherwise, the winning government would not be able to control the equilibrium fraction of entrepreneurs through policies.

Probabilistic voting has been used in similar static models with heterogeneous agents, endogenous credit constraints and occupational choice, where individuals vote based on an *ex-ante* position in society (*ex-ante* occupation) (see Fischer and Huerta, 2021; Huerta, 2023). This paper incorporates two additional challenges. First, agents vote before knowing their occupations, which means that the wealth distribution Γ_t is a state variable.[41] Thus, when voting, they anticipate which position in society they will occupy given some b_t. Secondly, even when new elections take place each period t, the equilibrium platform depends on the endogenous wealth distribution Γ_t, which is a function of the entire history of transfer rate policies.

B.1.3 Timing

The timing of events at period t is as follows: 1) Candidate A and B simultaneously and non-cooperatively announce their electoral platforms, b_t^A and b_t^B after observing Γ_t. 2) Elections are held, in which voters choose between the two candidates after observing their wealth a and Γ_t. 3) The elected candidate implements his announced transfer rate. 4) Given the winning platform b_t and wealth distribution Γ_t, banks define the minimum collateral required for a loan, $\hat{a}_t$. 5) After observing b_t and $\hat{a}_t$, agents make their occupational choice and their consumption-saving decisions.

B.1.4 The political objective function

I study the policy outcome when politicians maximize their share of votes. It is useful to start by identifying the 'swing voter' ($v_t = V_t$) for each level of wealth a. That is, the voter who is indifferent between the two candidates:

$$V_t = \bar{y}_t(a, b^A) - \bar{y}_t(a, b^B).$$

[41]This is a distinctive feature of my model. Typically, the papers that use probabilistic voting models take the different groups of agents (e.g. old/young, worker/entrepreneur) as given before elections take place. In contrast, in my setting these groups are endogenously defined after the equilibrium platform has been set. Interestingly, when voting, agents anticipate the implications of a given policy platform for occupational choice and choose their preferred candidate accordingly.

All agents with wealth a whose ideology is such that $v_t < V_t$ vote for candidate A, while the rest vote for B. Thus, the fraction of agents with wealth a from occupation j that vote for candidate A in period t is given by:

$$\pi_t^j(a, b^A) = Prob[v_t < V_t] = \phi^j \cdot [\bar{y}_t(a, b^A) - \bar{y}_t(a, b^B)] + \frac{1}{2}.$$

The share of votes of candidate A when it announces b^A, denoted by $S(b^A)$, is obtained by integrating $\pi_t^j(a, b^A)$ in wealth:

$$\begin{aligned} S(b^A) &= \int_{a<a_t^o(b^A)} \pi_t^W(a, b^A)\, d\Gamma_t(a) + \int_{a\geq a_t^o(b^A)} \pi_t^E(a, b^A)\, d\Gamma_t(a) + \frac{1}{2}, \\ &= \int_{a<a_t^o(b^A)} \phi^W \cdot [\bar{y}_t(a, b^A) - \bar{y}_t(a, b^B)] d\Gamma_t(a) + \int_{a\geq a_t^o(b^A)} \phi^E \cdot [\bar{y}_t(a, b^A) - \bar{y}_t(a, b^B)] d\Gamma_t(a) + \frac{1}{2}. \end{aligned}$$

Since both candidates maximize the their share of votes in period t, the Nash equilibrium is characterized by:

$$\begin{aligned} b_t^A &= \underset{b^A\in[\underline{b},\bar{b}]}{\arg\max} \left\{ \int_{a<a_t^o(b^A)} \phi^W \cdot [\bar{y}_t(a, b^A) - \bar{y}_t(a, b^B)] d\Gamma_t(a) + \int_{a\geq a_t^o(b^A)} \phi^E \cdot [\bar{y}_t(a, b^A) - \bar{y}_t(a, b^B)] d\Gamma_t(a) \right\}, \\ b_t^B &= \underset{b^B\in[\underline{b},\bar{b}]}{\arg\max} \left\{ \int_{a<a_t^o(b^B)} \phi^W \cdot [\bar{y}_t(a, b^B) - \bar{y}_t(a, b^A)] d\Gamma_t(a) + \int_{a\geq a_t^o(b^B)} \phi^E \cdot [\bar{y}_t(a, b^B) - \bar{y}_t(a, b^B)] d\Gamma_t(a) \right\}. \end{aligned}$$

The candidates' problems are symmetric. Thus, the policy platforms converge in equilibrium to the same transfer rate, b_t that maximizes the weighted income given the cumulative wealth distribution, Γ_t:

$$b_t = \underset{b\in[-\underline{b},\bar{b}]}{\arg\max} \left\{ \phi^W \int_{a<a_t^o(b)} \left(w_t(b)\ell + \frac{T_t(b) - \tau_t(b)a}{\phi^W} \right) d\Gamma_t(a) + \phi^E \int_{a\geq a_t^o(b^B)} \left(\Pi_t(b) - \frac{\tau_t(b)a}{\phi^E} \right) d\Gamma_t(a) \right\}. \tag{B.3}$$

Note that when choosing b candidates take into account the effects on a^o, which determines occupational choice and the demographic weights of workers and entrepreneurs.

Dividing by ϕ^W and defining the political weight $\phi \equiv \frac{\phi^E}{\phi^W} > 1$, the politicians' problem is written as follows:

$$b_t = \arg\max_{b \in [-\underline{b}, \bar{b}]} \left\{ \mathcal{W}(b, \Gamma_t) \equiv w_t(b)\, \Gamma(a_t^0(b)) + \phi\, \Pi_t(b)\, \left(1 - \Gamma(a_t^0(b))\right) \right\}. \tag{B.4}$$

Noting that the fraction of entrepreneurs is $e_t = 1 - \Gamma(a_t^0(b))$, the political objective function reads as $\mathcal{W}(b, \Gamma_t) = w_t \ell (1 - e_t) + \phi \Pi_t e_t$. Thus, the problem of politicians coincides with problem (5.3) presented in Section 5.

B.2 Constant risk aversion utility

Consider the CRRA utility function: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. As in the main text, assume that agents cannot anticipate the joint evolution of the transfer rate and the wealth distribution given some current transfer rate at t. Then, consumption-savings policy function are as in the following lemma.

Lemma 7 *The optimal consumption and savings policy functions that solve the HJB equation* (3.10) *are given by:*

$$c_t(a) = (1 - \theta_t) \cdot y_t(a),$$

$$s_t(a) = \theta_t \cdot y_t(a),$$

where the saving rate is: $\theta_t = \frac{1}{\sigma}\left(1 - \frac{\rho}{r - \tau_t}\right)$.

Proof:

Define the occupation-specific income as follows:

$$\tilde{y}(a) = (w\ell + T) \cdot \mathbb{1}_{a < a^o} + \Pi \cdot \mathbb{1}_{a \geq a^o}$$

Guess that the value function is $v(a) = B \cdot \frac{\left(a + \frac{\tilde{y}}{r-\tau}\right)^{1-\sigma}}{1-\sigma}$. Thus, $v'(a) = B\left(a + \frac{\tilde{y}}{r-\tau}\right)^{\sigma}$.

The FOC of (3.10) is:

$$u'(c) = v'(a) \Rightarrow c(a) = B^{\frac{1}{\sigma}} \left(a + \frac{\tilde{y}}{r(1-\tau)} \right).$$

Then the HJB equation reads as:

$$\begin{aligned}
\rho B \frac{\left(a + \frac{\tilde{y}}{r-\tau}\right)^{1-\sigma}}{1-\sigma} &= B^{-\frac{1-\sigma}{\sigma}} \frac{\left(a + \frac{\tilde{y}}{r-\tau}\right)^{1-\sigma}}{1-\sigma} + B\left(a + \frac{\tilde{y}}{r-\tau}\right)^{-\sigma} \left(\tilde{y} + (r-\tau)a - B^{\frac{-1}{\sigma}} \left(a + \frac{\tilde{y}}{r-\tau}\right)\right), \\
&= B^{-\frac{1-\sigma}{\sigma}} \frac{\left(a + \frac{\tilde{y}}{r-\tau}\right)^{1-\sigma}}{1-\sigma} + B\left((r-\tau) - B^{\frac{-1}{\sigma}}\right) \left(a + \frac{\tilde{y}}{r-\tau}\right)^{1-\sigma}, \\
\Leftrightarrow \frac{\rho B}{1-\sigma} &= \frac{B^{\frac{-(1-\sigma)}{\sigma}}}{1-\sigma} + B(r-\tau) - B^{\frac{-1}{\sigma}}.
\end{aligned}$$

Therefore, $B^{\frac{-1}{\sigma}} = \frac{\rho-(1-\sigma)(r-\tau)}{\sigma}$. Replacing in the FOC: $c(a) = \frac{\rho-(1-\sigma)(r-\tau)}{\sigma}\left(a + \frac{\tilde{y}}{(r-\tau)}\right)$. Finally, the individual budget constraint implies: $s(a) = \frac{(r-\tau)-\rho}{\sigma}\left(a + \frac{\tilde{y}}{r-\tau}\right)$. Using that $y = (r-\tau)a + \tilde{y}$ and defining $\theta = \frac{1}{\sigma}\left(1 - \frac{\rho}{r-\tau_t}\right)$ lead to the expressions in the lemma. ■

B.3 Transfers to workers and entrepreneurs

In this section , I extend the model to allow for a more realistic policy instrument in which candidate governments can provide positive transfers to both workers and entrepreneurs.

B.3.1 Government's budget

In each period, government spending is given by $G_t = g_t \cdot Y_t$, where $g_t \in [0, 1]$ is the fraction of GDP that is used to finance benefits to workers and entrepreneurs. The government chooses the share of total spending $b_t \in [\text{-}\underline{b}, \overline{b}]$ that finances transfers to workers. The rest is used to finance business policies. Total transfers to workers are given by $T_t^W = b_t \cdot G_t$, while transfers to entrepreneurs are $T_t^E = (1-b_t) \cdot G_t$.

The government finances transfers by levying a capital tax τ_t to all agents while keeping a balanced budget in each period:

$$\tau_t A_t = T_t^W(1-e_t) + T_t^E e_t.$$

B.3.2 Workers and entrepreneurs' budgets

The budget constraint of an agent with assets a_t is given by:

$$\dot{a}_t = (r-\tau_t)a_t - c_t + \begin{cases} w_t\ell + T_t^W & \text{if } a_t < a_t^o, \\ \Pi_t + T_t^E & \text{if } a_t \geq a_t^o, \end{cases}$$

where the occupational threshold a_t^o is defined in the next section.

B.3.3 Credit constraints and occupational choice

The financial market works as is the main text. The incentive compatibility (IC) constraint is given by:

$$\Pi_t + ra + T_t^E \geq (I-a) + w_t\ell + T_t^W.$$

Thus, the minimum collateral to obtain credit is implicitly defined by

$$\hat{a}_t = I - \frac{p_t R - w_t\ell - (T_t^W - T_t^E)}{1+r},$$

where $T_t^W - T_t^E = g_t \cdot Y_t(2b_t - 1)$ is increasing in the share of spending that finances social benefits (b_t). If $b_t \geq \frac{1}{2}$, the minimum collateral and prices preserve the same properties established in the paper. That is, when the transfer rate increases, $\hat{a}$ goes up, reducing the fraction of agents that can start a firm, and thus, increasing p but decreasing w.

In Section B.3.4, when I describe the political equilibrium, I show that having $b_t \geq \frac{1}{2}$ translates into some restrictions on the exogenous parameters of the model. In any case, the data supports this restriction. In Section D.2, I find that spending in social benefits as a fraction of GDP is much higher than that on business policies.

The occupational constraint is

$$\Pi_t + T_t^E \geq w_t \ell + T_t^W,$$

which defines a occupational threshold, $\tilde{a}_t$. As before, the effective occupational threshold that defines the first agent that starts a firm is given by: $a^o(b_t, g_t, \Gamma_t) \equiv \max\{\hat{a}_t, \tilde{a}_t\}$.

B.3.4 The government problem

In each period, a candidate government chooses the transfer rate to maximize the same objective function, $\mathcal{W}(b, g_t, \Gamma_t) = w_t \ell (1 - e_t) + \phi \Pi_t e_t$, while taking as given the share of GDP used to finance transfers g_t and the wealth distribution, Γ_t. As in the baseline model, the solution of the government problem satisfies:

$$1 - \Gamma_t \left(a^o(b_t, g_t, \Gamma_t) \right) = e_t^*,$$

where e_t^* is given by the fraction of entrepreneurs that solves equation (A.29).

Thus, the effective occupational threshold is $a_t^o = \Gamma_t^{-1}(1 - e_t^*)$. The equilibrium transfer rate is given by:

$$b_t = \begin{cases} \frac{(\varphi r I - 1)I + (1+r)a_t^o}{2 g_t Y_t^{PE}} + \frac{1}{2} & \text{if } a_t^o < \frac{I}{1+r}, \\ \frac{\varphi r I}{2 g_t Y_t^{PE}} + \frac{1}{2} & \text{if } a_t^o \geq \frac{I}{1+r}, \end{cases}$$

where $\varphi = \frac{(\phi-1)(1-\alpha)}{1-\alpha(1-\phi)}$ and the equilibrium output is given by: $Y_t^{PE} = Z_t (R e_t^*)^\alpha (\ell (1 - e_t^*))^{1-\alpha}$.

Therefore, the maximum sustainable transfer rate is $\overline{b} = \frac{\varphi r I}{2 g_t Y_t^{PE}} + \frac{1}{2}$. Imposing that $\overline{b} = 1$, gives an expression for the size of the government at period t: $g_t = \frac{\varphi r I}{Y_t^{PE}}$. Further, in order to satisfy that $b_t \geq \frac{1}{2}$ it must be that $\varphi r I \geq 1$, so that $\underline{b} \geq \frac{1}{2}$. This condition guarantees that the results regarding the transition dynamics presented in the main tex still hold.[42]

[42] A key step to characterize the transition dynamics is to show that the term is the square brackets in equations (A.16) and (A.19) is positive (the term I denote by χ_t). In the baseline model, Assumption 1 is sufficient to guarantee that $\chi_t > 0$. That condition holds trivially here, since $b_t \geq 0$. However, the form of χ_t changes and a sufficient condition for it to be positive is that $b_t \geq \frac{1}{2}$.

B.4 Capital and labor taxes

In this section, I extend the model to include both a labor and capital tax. To simplify the analysis, I follow Krusell and Rios-Rull (1999) and assume that capital and labor are taxed at the common rate, τ_t.

B.4.1 Government's budget

In each period, the government decides the share of total spending, $b_t \in [-\underline{b}, \overline{b}]$ to finance social benefits to workers. Then, it chooses the income tax τ_t to keep a balanced budget:

$$\tau_t A_t + \tau_t w_t (1 - e_t) = T(1 - e_t)$$

B.4.2 Workers and entrepreneurs' budgets

The individual budget constraint is given by:

$$\dot{a}_t = \begin{cases} (r - \tau_t) a_t + w_t \ell (1 - \tau_t) + T_t & \text{if } a_t < a_t^o, \\ (r - \tau_t) a_t + \Pi_t & \text{if } a_t \geq a_t^o, \end{cases} \tag{B.5}$$

B.4.3 Credit constraints and occupational choice

The IC constraint is given by

$$\Pi_t + ra \geq (I - a) + w_t \ell (1 - \tau_t) + T_t.$$

The minimum collateral is implicitly defined by

$$\hat{a}_t = I - \frac{p_t R - w_t \ell (1 - \tau_t) - T_t}{1 + r}.$$

The OC constraint is

$$\Pi_t \geq w_t \ell(1-\tau_t) + T_t.$$

which defines the occupational threshold, $\tilde{a}_t$. Note that an important difference with respect to the baseline model and the previous extension is that the occupational threshold $a_t^o = \max\{\hat{a}_t, \tilde{a}_t\}$ depends on the tax rate τ_t. This complicates the analysis. In particular, when b goes up, labor taxes increase, reducing to some extent the labor income. This effect relaxes the IC and OC constraints.[43] Therefore, the impact of b on the IC and OC constraints is ambiguous.

A crucial property for the main results to hold is that the occupational threshold is increasing in the transfer rate, b. The following lemma states some sufficient conditions for this property to be satisfied.

Lemma 8 *If $e_t \geq \frac{\alpha}{2}$ and $Y_t \leq \frac{1}{1-\alpha}$, then the occupational threshold a_t^o is increasing in b.*

Proof: Rewrite the budget constraint of the government as follows

$$\begin{aligned}
&\tau_t(A_t + (1-\alpha)Y_t) = bY_t(1-e_t),\\
\Rightarrow\ & d_b\tau_t(A_t + (1-\alpha)Y_t) = Y_t(1-e_t) + \underbrace{\left[b(1-e_t)\left(\frac{\alpha - 2e_t}{e_t(1-e_t)}\right) - \tau_t(1-\alpha)\frac{(\alpha - e_t)}{e_t(1-e_t)}\right]}_{\equiv \Omega_t} Y_t\, d_b e_t.
\end{aligned}$$

The term in the square brackets is negative as long as $e_t \geq \frac{\alpha}{2}$.[44] The IC constraint implies that

$$\begin{aligned}
d_b\hat{a}_t(1+r) &= \left[w\ell(1-\tau_t)\left(\alpha\left(\frac{1}{e_t} + \frac{1}{1-e_t}\right) - b\right) + pR\left((1-\alpha)\left(\frac{1}{e_t} + \frac{1}{1-e_t}\right) + b\right)\right] d_t e_t + Y_t - w_t \ell d_b \tau_t\\
&= \tilde{\chi}_t d_t e_t + Y_t(1 - Y_t(1-\alpha)) - w_t \ell \Omega_t Y_t d_b e_t\\
\Rightarrow d_b\hat{a}_t &= \frac{Y_t(1 - Y_t(1-\alpha))}{1 + r + (\tilde{\chi}_t + w_t \ell \Omega_t Y_t)\gamma(\hat{a}_t)}
\end{aligned}$$

[43] Recall that in the baseline model, when b increases there are three second order effects that relax the IC and OC constraints: i) the price of capital goes up, ii) the wage rate decreases, and iii) output decreases, reducing transfers to workers. In this extension, there is an additional second order effect through labor income taxes.

[44] Recall that at the political equilibrium, $e_t \leq \alpha$ (see Proposition 6).

Similarly, $d_b \tilde{a}_t = \frac{Y_t(1-Y_t(1-\alpha))}{(\tilde{\chi}_t + w_t \ell \Omega_t Y_t)\gamma(\hat{a}_t)}$. Thus, if $Y_t \leq \frac{1}{1-\alpha}$, then $a_t^o = \max\{\hat{a}_t, \tilde{a}_t\}$ is increasing in b. ■

B.4.4 The government problem

The government problem is the same as in the main text. Thus, the equilibrium transfer rate is implicitly given by: $1 - \Gamma_t(a^o(b,\Gamma_t)) = e_t^*$. Suppose that the aggregate productivity satisfies $Z_t \in [\underline{Z}, \overline{Z}]$. To satisfy the conditions of Lemma 8, it is sufficient to have that:

i) $e^*(\underline{Z}) \geq \frac{\alpha}{2}$, and ii) $\overline{Z}(Re^*(\overline{Z}))^\alpha(\ell(1-e^*(\overline{Z}))^{1-\alpha} \leq \frac{1}{1-\alpha}$.

Both conditions impose restrictions on the exogenous parameters of the model. If i) and ii) hold, then all the proofs still go through.

B.5 Forward-looking government

Consider a forward-looking government that in each period decides the transfer rate, b_t that maximizes the discounted sum of utilities given the wealth distribution Γ_t. Formally, the government solves:

$$\max_{b \geq -\underline{b}} \left\{ \mathcal{W}(b, \Gamma_t) \equiv \int_{a < a^o(b,\Gamma)} v_t(a) d\Gamma_t(a) + \int_{a \geq a^o(b,\Gamma)} v_t(a) d_t\Gamma(a) \right\}. \tag{B.6}$$

Lemma 9 provides an expression that implicitly defines the equilibrium transfer rate, b_t. Note, however, that the evolution of b_t depends on the dynamics of the entire wealth distribution. Thus, characterizing the transition dynamics as in the main text becomes analytically untractable and one must rely on numerical methods.

Lemma 9 *The transfer rate b that solves* (B.6) *satisfies:*

$$\int_{a < a^o(b,\Gamma_t)} \frac{(d_b w_t \ell + d_b T_t)}{y_t(a)} d\Gamma_t(a) + \int_{a \geq a^o(b,\Gamma_t)} \frac{d_b p_t}{y_t(a)} d\Gamma_t(a) = d_b \tau_t \int \frac{a}{y(a)} d\Gamma_t(a) + e^{\rho t} \left(\int_t^{+\infty} d_b \tau_s \frac{1}{r - \tau_s} e^{-\rho s} ds \right) + \frac{1}{\rho}. \tag{B.7}$$

Proof: From the proof of Lemma 1, recall that:

$$\rho v_t(a) = \log(y_t(a)) + \tilde{v}_t, \tag{B.8}$$

where

$$\tilde{v}_t = \log(\rho) - (\log(r - \tau_t) + \frac{r - \tau_t}{\rho} - 1 + \frac{1}{\rho} d_t \tilde{v}_t.$$

This last equation implies that:

$$\begin{aligned}\int_t^{+\infty} (\rho \tilde{v}_s e^{-\rho s} - d_s \tilde{v}_s e^{-\rho s}) ds = (\rho \log(\rho) - \rho) e^{-\rho t} + \int_t^{+\infty} (\log(r - \tau_s) + r - \tau_s) e^{-\rho s} ds, \\ \Rightarrow \tilde{v}_t = (\rho \log(\rho) - \rho) + e^{\rho t} \int_t^{+\infty} (\log(r - \tau_s) + r - \tau_s) e^{-\rho s} ds.\end{aligned}$$

Therefore,

$$d_b \tilde{v}_t = -e^{\rho t} \left(\int_t^{+\infty} \left(\frac{d_b \tau_s}{r - \tau_s} \right) e^{-\rho s} ds \right) - \frac{1}{\rho}.$$

Differentiating (B.8) in terms of b gives

$$d_b \rho v_t = d_b \tilde{v}_t + \begin{cases} \frac{d_b w_t \ell + d_b T_t - d_b \tau_t a}{y_t(a)} & \text{if } a < a^o(b, \Gamma_t), \\ \frac{d_b p_t - d_b \tau_t a}{y_t(a)} & \text{if } a \geq a^o(b, \Gamma_t). \end{cases}$$

Thus, the first order condition of $\max\{\rho \mathcal{W}(b, \Gamma_t)\}$ is given by

$$\int_{a < a^o(b,\Gamma_t)} \frac{(d_b w_t \ell + d_b T_t)}{y_t(a)} d\Gamma_t(a) + \int_{a \geq a^o(b,\Gamma_t)} \frac{d_b p_t}{y_t(a)} d\Gamma_t(a) - d_b \tau_t \int \frac{a}{y(a)} d\Gamma_t(a) - e^{\rho t} \left(\int_t^{+\infty} \left(\frac{d_b \tau_s}{r - \tau_s} \right) e^{-\rho s} ds \right) - \frac{1}{\rho} = 0.$$

Rearranging terms leads to equation (B.7). ■

B.6 First order stochastic dominance and the evolution of social benefits

This section is complementary to Section 6.1 as it shows what are the effects of applying an MIT shock to the wealth distribution according to First Order Stochastic Dominance (FOSD).

For the following lemma, define the set of initial distributions $\Gamma_0 \in \mathcal{G}_0$ that FOSD the stationary distribution $\Gamma^* \in \mathcal{G}$.

$$\mathcal{G}_{FOSD}(\Gamma^*) \equiv \{\Gamma_0 \,:\, \Gamma_0(a) \leq \Gamma^*(a), \forall a, \Gamma^* \in \mathcal{G}\}.$$

Similarly, define the set of initial distributions that reverse-first order stochastic dominate (reverse-FOSD) the steady-state distribution $\Gamma^* \in \mathcal{G}$ as:

$$\mathcal{G}_{FOSD_{rev}}(\Gamma^*) \equiv \{\Gamma_0 \,:\, \Gamma_0(a) \geq \Gamma^*(a), \forall a, \Gamma^* \in \mathcal{G}\}.$$

Restrict the FOSD distributions Γ_0 around Γ^* to those such that their mean A_0 belongs to some close neighbourhood around A^*, denoted by $A_0 \in N_\varepsilon(A^*)$.

Lemma 10 *Consider some stationary distribution* $\Gamma^* \in \mathcal{G}$.

1. *Take* $\Gamma_0 \in \mathcal{G}_{FOSD}(\Gamma^*)$ *with* $A_0 \in N_\varepsilon(A^*)$*, then: i)* $d_t b_t \leq 0$*, ii)* $d_t A_t \leq 0$*, iii)* $d_t a^o \leq 0$.

2. *Take* $\Gamma_0 \in \mathcal{G}_{FOSD_{rev}}(\Gamma^*)$ *with* $A_0 \in N_\varepsilon(A^*)$*, then the signs of i) to iv) are reversed.*

Proof: Consider the initial distribution $\Gamma_0 \in \mathcal{G}_{FOSD}(\Gamma^*)$. Thus, $d_t\Gamma_t(a)\big|_{t=0} \leq 0$. A similar procedure used in the proof of Proposition 3 shows that:

$$d_t b_t\big|_{t=0} = -\frac{d_t\Gamma(\hat{a}_t)\big|_{t=0}(1+r)}{d_a\Gamma_0(\hat{a}_0)Y_0} \geq 0.$$

The effect on τ is given by:

$$d_t\tau_t\big|_{t=0} = d_t b_t\big|_{t=0} Y_0(1-e_0) - \tau_0 d_t A_t\big|_{t=0}$$

Thus, $d_t\tau_t\big|_{t=0} \geq 0$ as long as $d_t A_t\big|_{t=0} \geq 0$ is small, i.e. if $A_0 \in N_\varepsilon(A^*)$. Hence, the tax rate goes up: $\tau_0 \geq r - \rho$. Proposition 2 implies that: i) $d_t b_t \leq 0$, ii) $d_t A_t \leq 0$, iii) $d_t a^o \leq 0$. The opposite holds if $\Gamma_0 \in \mathcal{G}_{FOSD_{rev}}(\Gamma^*)$. ■

To understand Lemma 10, consider an economy with some initial stationary wealth distribution. At period $t = 0$ the economy is hit by an MIT shock that shifts the distribution according to FOSD. The first order effect is that, given the previous transfer rate (b^*), the mass of entrepreneurs increases . Therefore, profits decrease, and thus, it becomes less attractive to start a firm. Thus, agents from the emerging class have diminished preferences for business-supporting policies, while the working class continues to demand high social benefits. As a result, social benefits increase at $t = 0$ ($b_0 > b^*$).

Additionally, at $t = 0$, aggregate wealth increases. If the mean effect is not too large compared to the distributive effect ($A_0 \in N_\epsilon(A^*)$), then the tax rate must increase at $t = 0$ to finance greater social benefits. Therefore, the cost of capital increases, causing agents to dissave over time. As a result, the wealth distribution shifts to the left in the FOSD sense. The shift in the distribution reduces the fraction of entrepreneurs, raising profits and increasing the demand for business-supporting policies from the emerging class. Consequently, social benefits decrease over time.

C Appendix: Additional Properties and Results

C.1 The occupational and incentive constraints

In this section, I show the properties of the IC and OC constraints illustrated in figure 4 of Section 3.1.1. In what follows, I take as given the transfer rate and the wealth distribution. Thus, I omit the dependence on (b, Γ). Recall that

$$OC(e) = h(e) - rI,$$

$$IC(e) = h(e) - x(e),$$

where

$$h(e) = p(e)R - w(e)\ell - bY(e), \tag{C.1}$$

$$x(e) = (1 + r)[I - \Gamma^{-1}(1 - e)]. \tag{C.2}$$

First, note that:

$\lim_{e\to 0} p(e) = +\infty, \lim_{e\to 0} w(e) = 0, \lim_{e\to 0} y(e) = 0,$

$\lim_{e\to 1} p(e) = 0, \lim_{e\to 1} w(e) = +\infty, \lim_{e\to 1} y(e) = 0.$

Thus, $\lim_{e\to 0} h(e) = +\infty$ and $\lim_{e\to 1} h(e) = -\infty$. Differentiating equation (C.1) gives

$$h'(e) = -\left(\frac{1}{e} + \frac{1}{1-e}\right)\big(\alpha w\ell + (1-\alpha)pR + b(pR - w\ell)\big) = -\chi < 0,$$

where I have used the fact that under Assumption 1, $\chi > 0$. Further, $\lim_{e\to 0} h'(e) = +\infty$, while $\lim_{e\to 1} h'(e) = -\infty$.

Secondly, note that:

$\lim_{e\to 0} x(e) = (1 + r)[I - \Gamma^{-1}(1)] = -\infty$ and $\lim_{e\to 1} x(e) = (1 + r)[I - \Gamma^{-1}(0)] = (1 + r)(I - \underline{a}).$

Further, $x'(e) = \frac{(1-r)e}{\gamma(a^o)} > 0$, where I have used that $d_e a^o = -\frac{e}{\gamma(a^o)}$. Because $h'(e) < 0$, $x'(e) > 0$,

$h(0) = +\infty$, $x(0) = -\infty$, $h(1) = -\infty$, $x(1) > 0$, and both functions are continuous in e, there exists a unique $\hat{e}$ such that the IC constraint binds: $IC(\hat{e}) = 0$. Moreover, the properties of $h(e)$ imply that there is a unique $\tilde{e}$ such that the OC constraint binds: $OC(\hat{e}) = 0$. Finally, both constraints intersect at a unique point e' that satisfies: $I - (1 + r)\Gamma^{-1}(e') = 0$. Therefore, at this point it must be that $a^o = \Gamma^{-1}(e') = \frac{I}{1+r}$. This completes the proof of the properties illustrated in Figure 4.

C.2 The competitive equilibrium according to the OC-IC diagram

Figure 13 illustrates the dynamics of the competitive equilibrium according to the OC-IC diagram and for Case 1 in Proposition 1. I consider three points of time: initial ($t = 0$), medium-run ($t = mr$), and long-run (lr). The economy starts from a wealth distribution Γ_0 such that $\tau_0 < r - \rho$.

First, from equation (C.1), note that $h(e)$ does not depend on the wealth distribution as shown by the gray line in the graph. Secondly, because the economy accumulates wealth, Γ shifts right over time. Therefore, $x(e)$ shifts right over time as illustrated by the dashed and dotted-dashed lines. Overall, the economy moves along the $h(e)$ curve over time, as indicated by the small arrows in the figure.

Initially, the IC constraint binds. Thus, the initial fraction of entrepreneurs, $e_0 = e(\Gamma_0)$, is given by the intersection of the gray ($h(e)$) and solid black ($x_0(e)$) lines. In the medium-run, $x(e)$ shifts right, but the collateral constraint is still binding as shown by the intersection of $h(e)$ and the dashed line ($x_{mr}(e)$). The fraction of entrepreneurs increases in the medium run, so $e_{mr} = e(\Gamma_{mr}) > e_0$. In the long-run, the economy hits the OC constraint as shown by the cross between $h(e)$, the dotted-dashed line ($x_{lr}(e)$), and the horizontal line at rI. The economy reaches the steady-state fraction of entrepreneurs e_{lr} at which the OC and IC intersect.

In Case 1 of Proposition 1, the tax rate is equal to $r - \rho$ in the long-run, which keeps the wealth distribution unchanged over time. Hence, $x(e)$ does not change from that point onwards as shown in the figure. However, in Cases 2 and 3, the wealth distribution keeps shifting right over time. Therefore, $x(e)$ continues moving right over time, i.e. the IC becomes non-binding. However, the steady-state fraction of entrepreneurs in both cases is still given by e_{lr} which makes the OC constraint to bind.

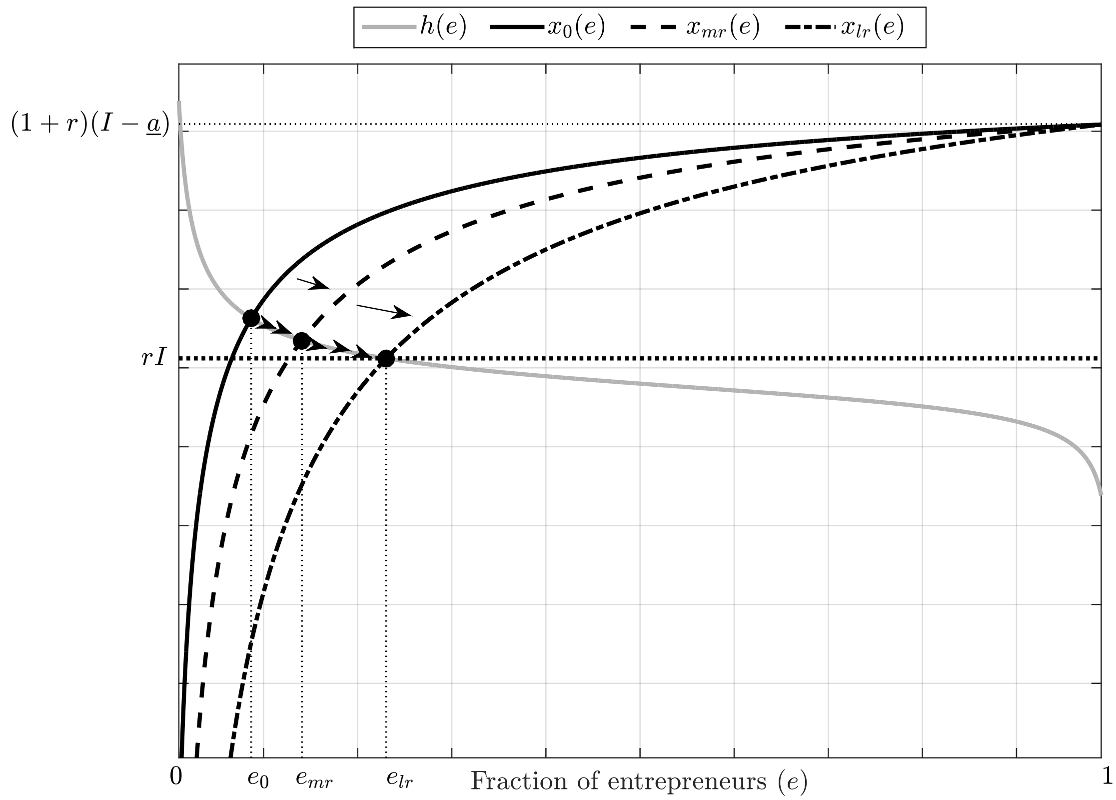


Figure 13: Competitive equilibrium: the OC-IC diagram.

C.3 A shift of the wealth distribution

In what follows, I consider a discrete time model where the length of a period is given by $\Delta > 0$. I illustrate the shift in the wealth distribution when agents save a positive fraction of their income at t, i.e. $\theta_t > 0$. The main conclusion is that Γ always shifts in the FOSD sense.

The saving policy function (3.8) can be written as follows:

$$s_t(a) = \begin{cases} \theta_t(r-\tau_t)a + \theta_t(w_t\ell + T_t) & \text{if } a < a_t^o, \\ \theta_t(r-\tau_t)a + \theta_t\Pi_t & \text{if } a \geq a_t^o. \end{cases}$$

Define $\underline{\omega}_t \equiv \theta_t(w_t \ell + T_t)$ and $\overline{\omega}_t \equiv \theta_t \Pi_t$. Then:

$$
\begin{aligned}
a_{t+\Delta} &= a_t + \Delta s_t(a_t), \\
&= a_t(1 + \Delta\theta(r - \tau_t)) + \Delta\underline{\omega}_t \mathbb{1}_{a_t < a_t^o} + \overline{\omega}_t \mathbb{1}_{a_t \geq a_t^o}.
\end{aligned} \tag{C.3}
$$

Consider the density function at period t, $\gamma_t(a)$ with support in $[\underline{a}_t, +\infty)$. Then, the density function at period $t + \Delta$, $\gamma_{t+\Delta}(a)$ is defined by:

$$
\gamma_{t+\Delta}(a) = \begin{cases}
0 & \text{if } a < \underline{a}_{t+\Delta}, \\
\gamma_t\left(\frac{a - \Delta\underline{\omega}_t}{1+\Delta\theta_t(r-\tau_t)}\right) & \text{if } a \in [\underline{a}_{t+\Delta}, a^1_{t+\Delta}), \\
0 & \text{if } a \in [a^1_{t+\Delta}, a^2_{t+\Delta}), \\
\gamma_t\left(\frac{a - \Delta\overline{\omega}_t}{1+\Delta\theta_t(r-\tau_t)}\right) & \text{if } a \geq a^2_{t+\Delta},
\end{cases} \tag{C.4}
$$

where I have defined:

$$
\begin{aligned}
\underline{a}_{t+\Delta} &\equiv \underline{a}_t(1 + \Delta\theta(r - \tau_t)) + \Delta\underline{\omega}_t, \\
a^1_{t+\Delta} &\equiv a^o_t(1 + \Delta\theta(r - \tau_t)) + \Delta\underline{\omega}_t, \\
a^2_{t+\Delta} &\equiv a^o_t(1 + \Delta\theta(r - \tau_t)) + \Delta\overline{\omega}_t.
\end{aligned}
$$

From equation (C.4), the next period cumulative wealth distribution reads as:

$$
\Gamma_{t+\Delta}(a) = \begin{cases}
0 & \text{if } a < \underline{a}_{t+\Delta}, \\
\Gamma_t\left(\frac{a - \Delta\underline{\omega}_t}{1+\Delta\theta_t(r-\tau_t)}\right) & \text{if } a \in [\underline{a}_{t+\Delta}, a^1_{t+\Delta}), \\
\Gamma_t(a^o_t) & \text{if } a \in [a^1_{t+\Delta}, a^2_{t+\Delta}), \\
\Gamma_t\left(\frac{a - \Delta\overline{\omega}_t}{1+\Delta\theta_t(r-\tau_t)}\right) & \text{if } a \geq a^2_{t+\Delta}.
\end{cases} \tag{C.5}
$$

Figures 24 and 25 in Section F depict the shift of the density and cumulative wealth distributions, respectively.

C.4 More on individual preferences

In this section, I dig deeper into individual preferences for redistribution. I explain how the preferred transfer rate function presented in Section 5.2 can be understood through the lens of the disposable income function.

First, denote by $b_W(a)$ the transfer rate that maximizes the income of a worker with asset a given some wealth distribution Γ. Similarly, denote by $b_E(a)$ the rate that maximizes the income of an entrepreneur. Secondly, denote by $\tilde{b}(a)$ the transfer rate that satisfies: $\tilde{b} : a = \hat{a}(\tilde{b})$. That is, the transfer rate that makes the minimum collateral exactly equal to the agent's assets. Figure 15 depicts $\tilde{b}$ as a function of assets. Thirdly, consider three agents with assets: $a_{WC} < a_{EC} < a_{IC}$, that belong to the working, emerging, and incumbent class, respectively.

The disposable income as a function of the transfer rate b for each level of assets is depicted in Figure 14. The solid line represents an agent from the working class ($a = a_{WC}$). This agent is not able to start a firm even under the most favorable policies for businesses (b = -$\underline{b}$), i.e. $a_{WC} < \hat{a}$(-$\underline{b}$). Thus, her preferred policy is given by $b^W(a_{WC})$, as indicated by the black dot in the figure. Wealthier agents from the working class must finance a larger fraction of social benefits through taxes. Therefore, the maximum of the workers' income shifts left as wealth increases. This explains why the preferred transfer rate is decreasing in assets within the working class.

The dotted line corresponds to an agent from the emerging class. This agent can start a firm only if $b \leq \tilde{b}$, for higher transfer rates she does not obtain credit and must become a worker. Thus, her disposable income function has a discrete fall at $b \leq \tilde{b}$. The upper curve corresponds to her income if she becomes an entrepreneur, while the lower curve represents her income when she becomes a worker. The disposable income function has two peaks, as indicated by the gray dot and the black dot to the right.

The most preferred transfer rate is represented by the gray dot ($b^E(a_{EC})$) that maximizes the agent's entrepreneurial income. However, this individual is constrained by the minimum collateral and cannot attain the first-best. Thus, she chooses $\tilde{b}(a_{EC})$ which is the maximum transfer rate that allows her to start a firm. At this transfer rate, the marginal benefit from increasing b is positive due to the price effect. That is, a higher transfer rate increases the minimum collateral, reducing competition and increasing prices. Because $\tilde{b}(a)$ is increasing in a, wealthier agents from the emerging class can afford to have a higher b to keep competition low. Therefore, the preferred transfer rate function is increasing in assets for collateral constrained agents. Sufficiently wealthy agents from the emerging class are not constrained by $\hat{a}$ and can choose the rate that maximizes their entrepreneurial income (the gray dot in the figure, $b^E(a_{EC})$).

Finally, agents from the incumbent class (dashed line) can always start a firm and choose $b^E(a_{IC})$, as indicated by the black dot in the figure. The maximum of their income shifts lefts with assets, because wealthier agents suffer more from higher capital taxes to finance social benefits. Thus, the preferred transfer rate is decreasing in assets within the incumbent class.

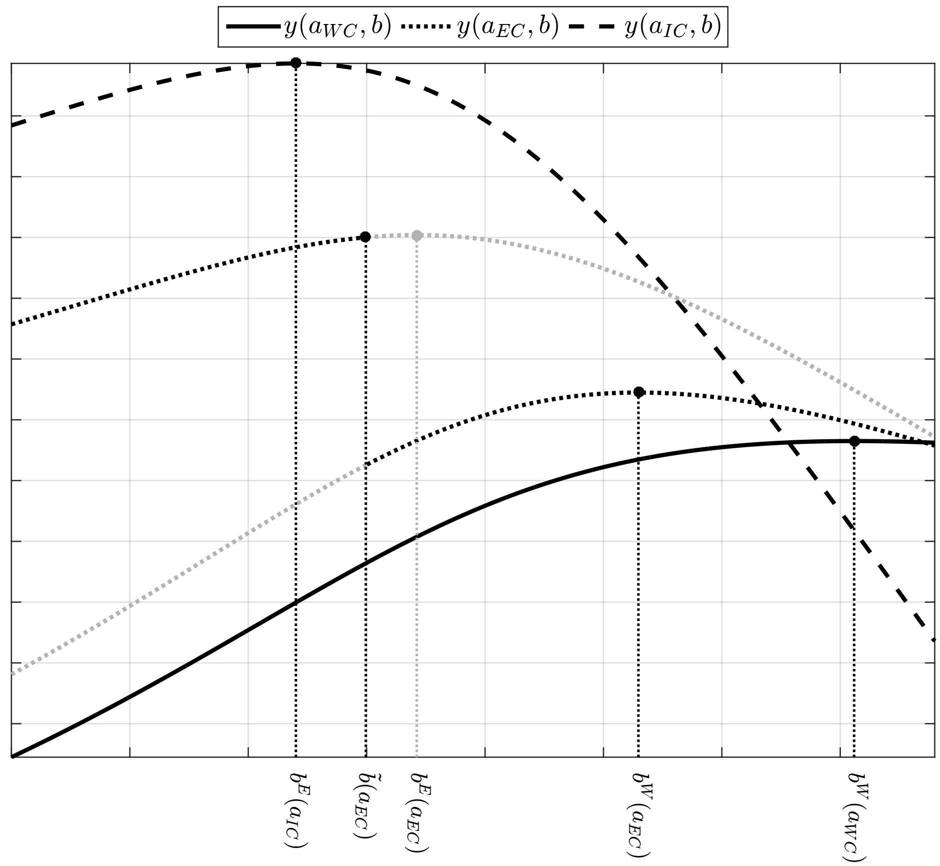


Figure 14: Disposable income as a function of assets and the transfer rate, $y(a, b)$.

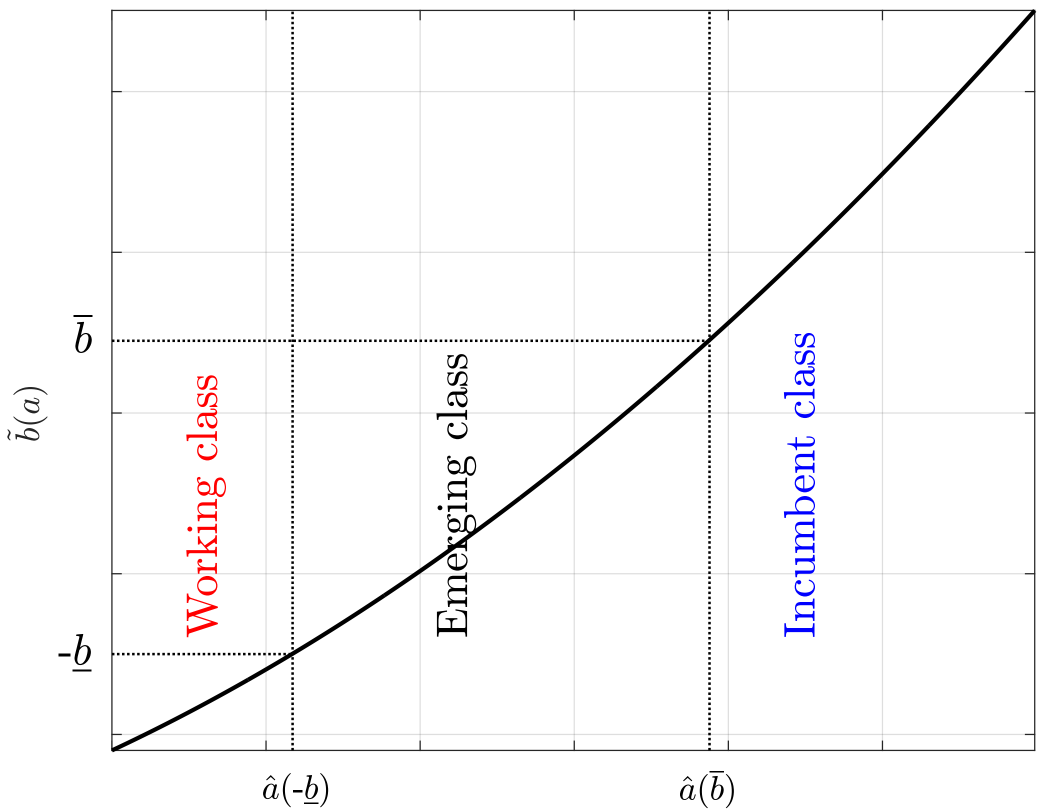


Figure 15: Maximum transfer rate to become an entrepreneur, $\tilde{b}(a)$.

C.5 The equilibrium size of the welfare state and the wealth distribution

Figure 16 illustrates equation (5.7) in presented in Section 5.3 that links the cumulative wealth distribution to the equilibrium size of the welfare state:

$$b_t = \begin{cases} \frac{(\varphi r-1)I+(1+r)a_t^o}{Y_{PE}} & \text{if } a_t^o \leq \frac{I}{1+r}, \\ \overline{b} \equiv \frac{\varphi r I}{Y_{PE}} & \text{if } a_t^o > \frac{I}{1+r}. \end{cases}$$

In order to express this relationship in a two dimensional graph, I invert the PE condition (5.4) to express the effective occupational threshold as follows:

$$a_t^o = \Gamma_t^{-1}(1 - e^*) \tag{C.6}$$

When the wealth distribution is such that $a_t^0 < \frac{I}{(1+r)}$, the IC constraint binds, and thus, the equilibrium size of the welfare state is represented by the black solid line in Figure 16. Otherwise, when $a_t^0 \geq \frac{I}{(1+r)}$, the OC condition binds and the size of the welfare state remains limited by the maximum sustainable transfer rate, $\overline{b}$ (dashed black line).

The blue and red arrows illustrate the effects on the transfer rate when the wealth distribution shifts right or left in the FOSD sense, respectively. According to equation (C.6), a right shift in Γ_t ($d_t\Gamma_t < 0$) increases the effective occupational threshold raising the transfer rate. Thus, the economy moves along the solid line and upwards.

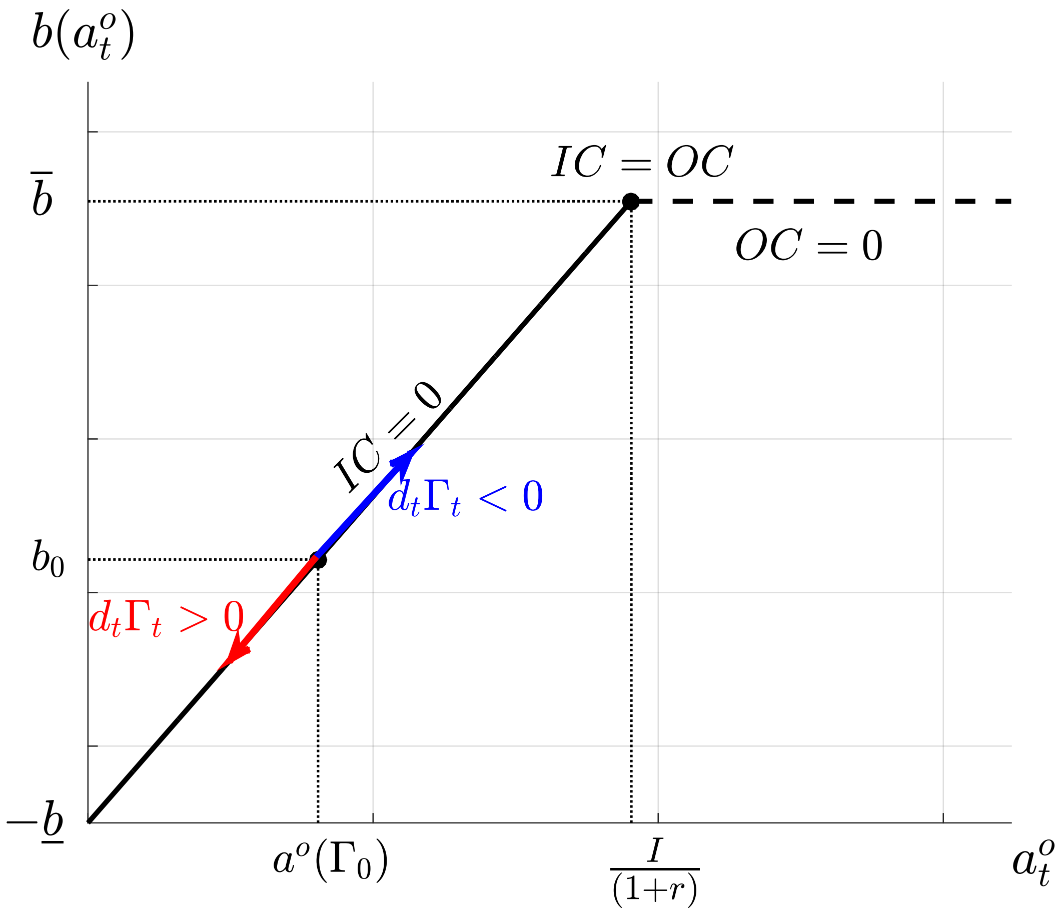


Figure 16: The PE condition.

C.6 The evolution of the welfare state according to the OC-IC diagram

Figure 17 illustrates the dynamics of the the welfare state according to the OC-IC diagram and for Case 1.(b) in Proposition 2. I consider three points of time: initial ($t = 0$), medium-run ($t = mr$), and long-run (lr). The economy starts from a wealth distribution Γ_0 such that $\tau_0 < r - \rho$. Additionally,the aggregate productivity is fixed over time. Thus, the fraction of entrepreneurs is given by e^*, while output is denoted by Y_{PE}.

The OC and IC constraints at the political equilibrium can be written as

$$OC(b) = \bar{b} - b, \tag{C.7}$$

$$IC_t(b) = \bar{b} - b - \kappa_t, \tag{C.8}$$

where recall that $\bar{b} = \frac{\varphi r I}{Y_{PE}}$ and I have defined $\kappa_t \equiv I - (1+r)\Gamma_t^{-1}(1-e^*)$. Note that the OC function (C.7) does not depend on time, while the IC function (C.8) depends on time only through the wealth distribution as captured by κ_t.

The gray horizontal line corresponds to the maximum sustainable transfer rate, $\bar{b}$. The solid black line is the 45° line. Thus, the difference between the gray and black line corresponds to the OC constraint function. The remaining black lines represent $b + \kappa_t$ at different points of time. Because initially $\tau_0 < r - \rho$, agents save and so the wealth distribution shifts right over time. Hence, $b + \kappa_t$ moves to the right in the FOSD sense over time.

The difference between the gray and dotted line is the IC constraint at $t = 0$. Their intersection gives the initial transfer rate, b_0. The economy accumulates wealth over time, and thus, the IC constraint shifts right (dashed line) increasing the transfer rate in the medium-run to b_{mr}. In Case 1(b).i of Proposition 2, the IC constraint remains binding in the long-run. Hence, the economy attains a steady-state transfer rate such that $b_{lr} < \bar{b}$. Overall, social benefits exhibit an increasing path over time as shown by the black arrows in the figure.

There is an additional case, 1(b).ii, in which the economy accumulates wealth indefinitely and so the transfer keeps increasing over time as indicated by the gray arrows. Eventually, the OC constraint becomes binding and the economy reaches the maximum sustainable transfer rate in the long-run, $\bar{b}$ (gray square in the figure).

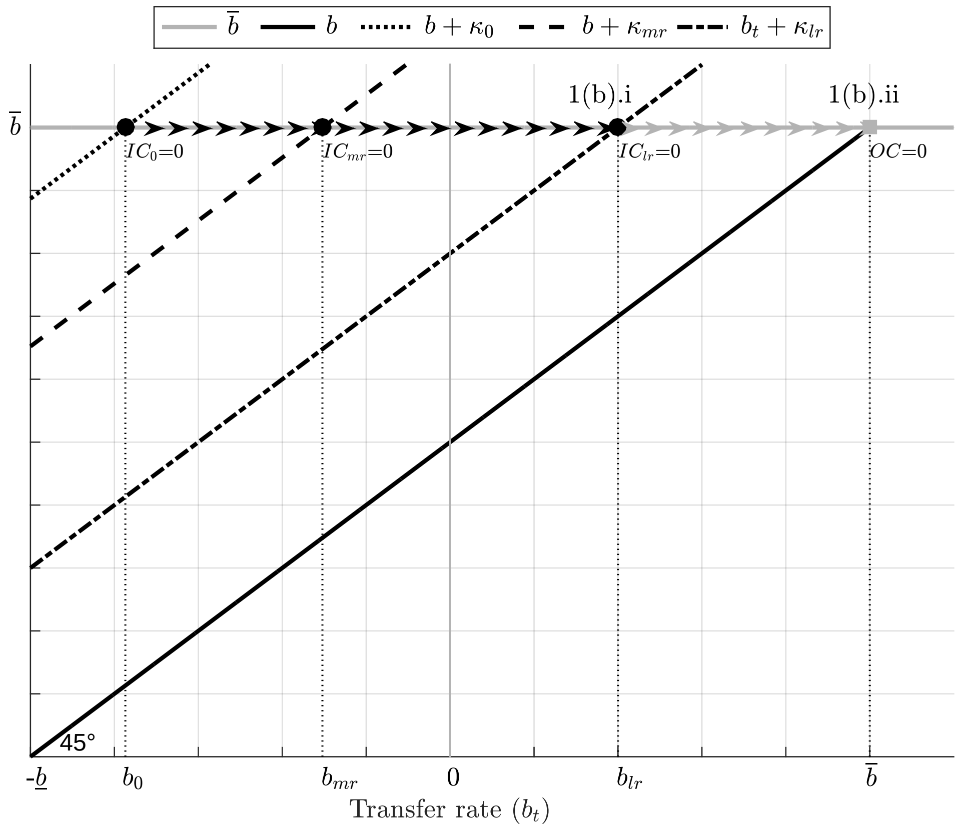


Figure 17: Political equilibrium: the IC-OC diagram.

C.7 The oscillatory behavior of the tax rate

In Section 6, I study the transition dynamics of the political equilibrium. There are six possible patterns. In particular, in Case 1(b).ii in Figure 8, the economy does not converge to a stationary distribution in the long-run. The explanation is that the economy reaches the maximum sustainable transfer rate (OC binds) before the tax rate can attain its stationary level, $r - \rho$. The tax rate converges to zero in the long-run ($\theta^* > 0$), and thus, the economy continues to accumulate wealth indefinitely.

There is another feature of the model that can prevent the economy to attain a stationary distribution: the "oscillatory behavior" of the tax rate. Consider Case 1(b) in Figure 8. Along the transition path, the transfer rate is increasing and the economy accumulates wealth. Thus, from the budget constraint of the government (3.14), the dynamics of the tax rate are given by:

$$d_t\tau_t = \frac{1}{A_t}\left(\underbrace{d_t b_t}_{>0} Y_{PE}(1-e^*) - \tau_t \underbrace{d_t A_t}_{>0}\right).$$

Hence, when $\tau_t < 0$, the tax rate increases over time. However, when $\tau_t > 0$, the sign of $d_t\tau_t$ is ambiguous. Thus, the tax rate can alternate between periods of growth and decline. This oscillatory behavior implies that τ may never reach its stationary level in the long-run. As a result, the economy may not attain a steady-state distribution as in case in Case 1(b).ii in Figure 8.

C.8 An aggregate productivity shock

In this section, I study the transition dynamics after an unexpected aggregate productivity shock. The economy is initially at some steady-state (b^*, Γ^*), such that $b^* > 0$ and the IC constraint binds ($a^o = \hat{a}^*$). Aggregate productivity is initially given by Z. At $t = 0$, there is a one-time marginal increase in aggregate productivity, $Z_0 > Z$. Thus, $Z_t = Z$ for $t > 0$. I assume that Z_0 satisfies: $Y_0 \equiv Y_{PE}(Z_0) \leq 1$.

The following lemma describes the transition dynamics after the aggregate productivity shock.

Lemma 11 *Consider an economy that is initially in a steady-state (b^*, Γ^*) with $b^* > 0$. At $t = 0$, there is a one-time marginal increase in aggregate productivity. First, the equilibrium transfer rate at $t = 0$ satisfies: $b_0 < b^*$. Further, $\tau_0 < r - \rho$ and $d_t A_0 > 0$.*

Secondly, denote by b_{0^+} the transfer rate right after the shock and by $b^{*\prime}$ the new stationary transfer rate. The dynamics of the political equilibrium for $t > 0$ are as follows:

1. *If $\tau_{0^+} < r - \rho$, then $d_t b_t \geq 0$, $d_t A_t \geq 0$, $d_t a_t^o \geq 0$, and $b^{*\prime} > b^*$.*

2. *If $\tau_{0^+} > r - \rho$, these effects are reversed.*

Proof: Use equation (A.28) to write the minimum collateral condition (2.6) as follows:

$$(1+r)\hat{a}_t = I + rI\left(1 - \frac{\phi}{1-\alpha(1-\phi)}\right) + b_t Y_t$$

Differentiating in terms of Z gives:

$$d_Z \hat{a}_0 = \frac{d_Z b_0 Y_0 + b_0 d_Z Y_0}{1+r}$$

The political equilibrium condition implies:

$$-\gamma_0(\hat{a}_0) d_Z \hat{a}_0 = d_Z e_0^*$$

Combining both conditions,

$$d_Z b_0 = -\frac{1+r}{\gamma_0(\hat{a}_0)Y_0} d_Z e_0^* - b_0 d_Z Y_0 < 0$$

where I have used that $d_Z Y_0 > 0$ and that $d_Z e_0^* > 0$ from equation (A.29). The tax rate satisfies:

$$d_Z \tau_0 = -\frac{1-e_0^*}{A^*}\left[\left(\frac{1+r}{\gamma_0(\hat{a}_0)Y_0} + \frac{b_0 Y_0}{1-e_0^*}\right) d_Z e_0^* + b_0(d_Z Y_0)(1-Y_0)\right] < 0,$$

where I have used that $Y_0 \leq 1$. Because the tax rate decreases at $t=0$, then $\theta_0 > 0$. Now I proceed to study the effects at $t = 0^+$. The political equilibrium condition (A.29) implies:

$$d_{t=0^+}\Gamma_t(\hat{a}_{0^+}) + \gamma_{0^+}(\hat{a}_{0^+}) d_{t=0^+}\hat{a}_t = d_{t=0^+} e_t^*$$

Using that $d_{t=0^+}\hat{a}_t = \frac{(d_{t=0^+} b_t) Y_{0^+} + b_{0^+}(d_{t=0^+} Y_t)}{1+r}$ leads to

$$d_{t=0^+} b_t = -\frac{1+r}{\gamma_{0^+}(\hat{a}_{0^+})}\left(d_{t=0^+} e_t^* + d_{t=0^+}\Gamma_t(\hat{a}_{0^+})\right) - b_{0^+}(d_{t=0^+} Y_t).$$

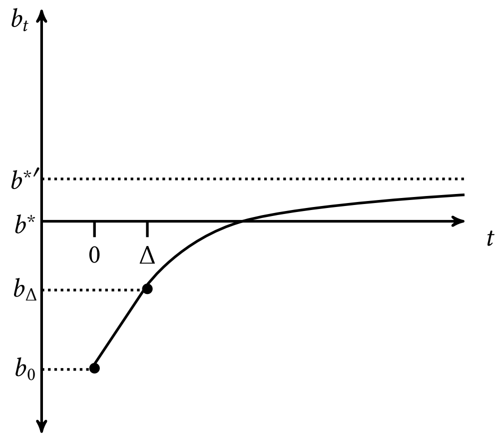
Because $\theta_0 > 0$, $d_{t=0^+}\Gamma_t(\hat{a}_{0^+}) \leq 0$. Also, $d_{t=0^+} Z_t < 0$, thus $d_{t=0^+} e_t^* < 0$ and $d_{t=0^+} Y_t < 0$. Hence,

$d_{t=0^+}b_t > 0$. Since $d_{t=0^+}b_t > 0$ and $d_{t=0^+}1 - e_t^* > 0$, and $d_{t=0^+}Y_t < 0$ and $d_{t=0^+}A_t > 0$, the effect on $\tau_{t=0^+}$ is ambiguous. If $\tau_{t=0^+} < r - \rho$, then Proposition 2 implies that $d_t b_t \geq 0$, $d_t A_t \geq 0$, $d_t a_t^o \geq 0$. Further, because the economy accumulates wealth, $b^{*\prime} > b^*$. Otherwise, these effects are reversed. ■

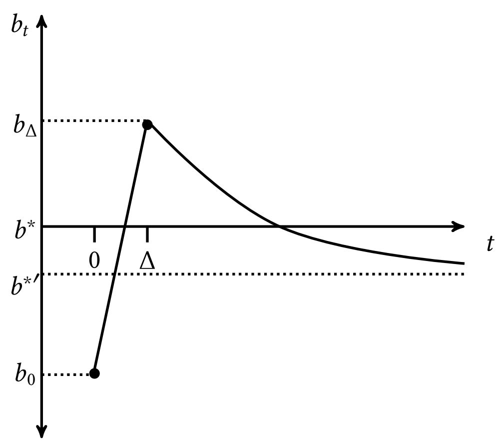
To understand Lemma 11, consider a discrete time model where the length of a period is $\Delta > 0$. Figure 18 illustrates the results in the lemma. Initially, at $t = 0$, a positive aggregate productivity shock increases the attractiveness of starting a business. Thus, the emerging class advocates for more business-supporting policies. As a result, social benefits initially decrease. The tax rate goes down, and thus, agents save and the wealth distribution shifts right in the FOSD sense.

In the next period, $t = \Delta$, there are three forces that push social benefits up. First, aggregate productivity goes down, so firms profits decrease. Secondly, because the distribution shifts right, more agents can start a firm, decreasing even more profits. Both effects reduce the attractiveness of starting a firm, so agents on average prefer higher social benefits. Thirdly, the aggregate productivity goes back to its initially lower value. Thus, total output decreases and the tax rate required to finance a certain amount of benefits increases. If the sum of these effects is too high, the cost of capital increases significantly. Hence, agents dissave and social benefits decrease over time, as shown by Case 2. in the figure. Otherwise, social benefits exhibit an increasing path as shown by Case 1.

Overall, a positive aggregate shock has an ambiguous effect on the evolution of social benefits. There are mainly two forces at play. The initial first order effect of having higher productivity that reduces social benefits, and the subsequent shift of the wealth distribution that causes social benefits to increase. In general, which effects dominates depends on the exogenous parameters and the properties of the wealth distribution.



(a) Case 1.



(b) Case 2.

Figure 18: Transition dynamics: an aggregate productivity shock.

D Appendix: Data

D.1 The evolution of social benefits across countries

In this section, I present data on the evolution of social benefits (SBs). I use the “social benefits to households" indicator available at the OECD National Accounts and Statistics database. This source offers “reliable and internationally comparable statistics on public and (mandatory and voluntary) private social expenditure at programme level".

SBs include transfers to households to provide support during circumstances which adversely affect their welfare. The National Accounts classify social benefits into two categories. The first category consists of SBs other than social transfers in-kind, which are typically provided in cash. Thus, households can use these benefits indistinguishably from other income sources. Examples include pensions, unemployment compensation, and maternity leave. The second category corresponds to the provision of goods and services, so households have no discretion over their usage. This includes areas such as education, health, housing assistance, and residential care.

The database includes information for the 38 OECD member countries and for 5 non-OECD countries. The time series extend until 2020, but I exclude this last year to rule out the effects of the pandemic. I select countries with available data starting from 1995 or earlier, ensuring a minimum of 25 years of data. This results in a sample of 31 OECD countries and 2 non-OECD countries.

The countries considered are: Australia, Austria, Belgium, Canada, Czech Republic, Switzerland, China, Germany, Denmark, Spain, Estonia, Finland, France, Great Britain, Greece, Hungary, Ireland, Israel, Italy, South Korea, Lithuania, Luxembourg, Latvia, Netherlands, Norway, New Zealand, Poland, Portugal, Slovakia, Slovenia, Sweden, United States, and South Africa. I classify these countries into 10 regions: North America, British Isles, Baltic States, Benelux, DACH (German-speaking European countries), Mediterranean, Scandinavia, Visegrád plus Slovenia, Non-Western countries, and Oceania.

Figure 19 presents the evolution of SBs as a fraction of GDP in the countries of each region.

In each figure, the black solid line corresponds to the total spending in SBs (in-cash plus in-kind transfers), while the dotted line represents cash transfers alone. The gray line is the trend of total SBs.[45]

Overall, the data indicates that there are large differences in the evolution of the welfare state across countries. 55% of the countries considered have been increasing their spending in SBs, while 21% have been decreasing spending in SBs. The remaining 24% of the countries have exhibited a relatively flat trend over time.

There are also important differences in the evolution of the welfare state within regions. In the United States, SBs have shown significant growth since the 1970's, while Canada has experienced a relatively stable trend. Among the DACH countries, Germany has seen a decline in SBs over the past 25 years, Switzerland has exhibited a slight increasing trend, and Austria has maintained a relatively flat trend. Scandinavia is perhaps the most surprising region, with SBs increasing in Finland and Norway, having a flat trend in Denmark, and consistently decreasing over time in Sweden. Lastly, Australia has seen an increase in SBs, while New Zealand has experienced a decline.

The data presented in Figure 19 corresponds to gross spending in social benefits. However, governments also provide social support through the tax system. The value of such tax measures is partially offset by the taxes paid by the recipients of social benefits. Therefore, net spending in social benefits is in general lower than gross spending. Figure 27 in section F presents the evolution of net social spending in the 10 regions considered. The data is also available at the OECD.

Many countries exhibit similar trends in net social spending compared to gross spending in social benefits, as observed in the United States and Sweden. However, there are countries, such as Greece and Netherlands, that experience a reversal in their trends when accounting for taxes. Overall, there are large differences in the evolution of net social spending. 60% of the countries

[45]China does not have data on spending on in-kind SBs. Thus, I use the spending on in-cash SBs instead. In general, cash transfers follow a similar pattern to that of total social benefits. Because the main focus is on the trend of SBs, using cash transfers should serve as a satisfactory proxy for the evolution of SBs in this country.

have been increasing net social spending, 23% have experienced a decrease, and the remaining 17% have maintained relatively stable levels.

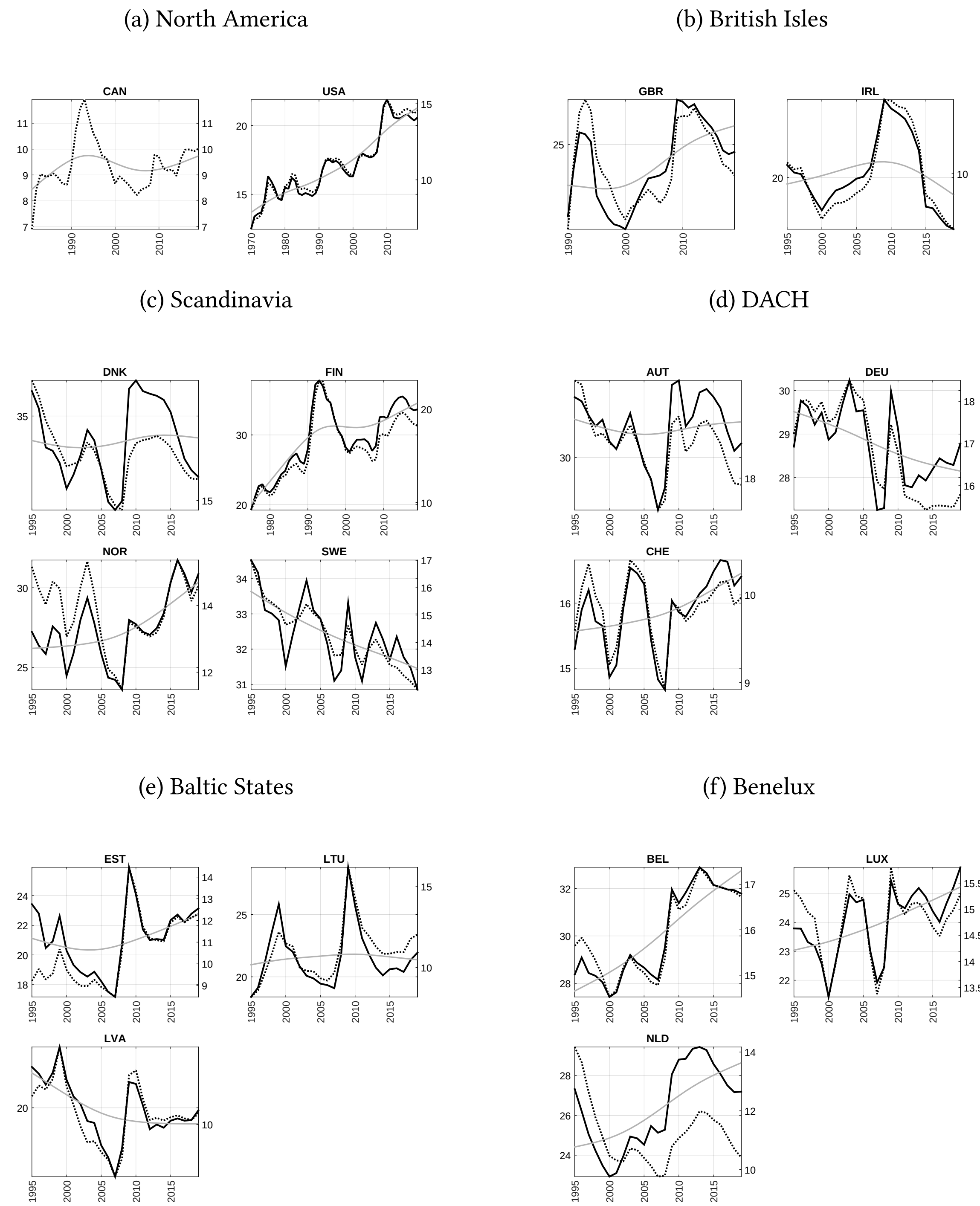
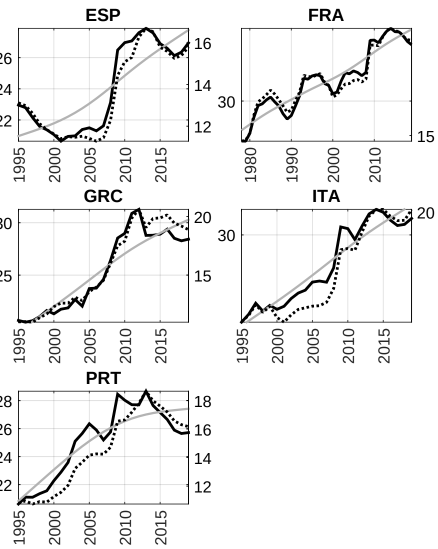


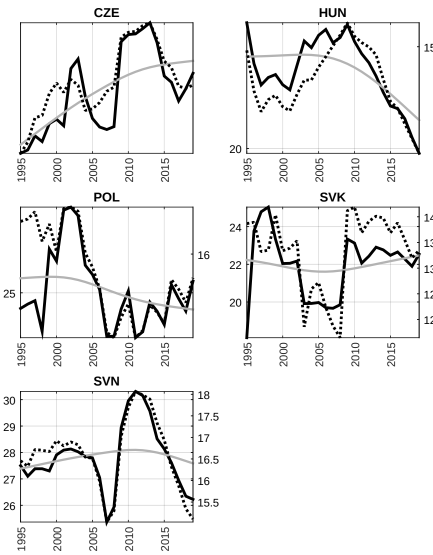
Figure 19: Social Benefits (% GDP).

Total: solid line (left y-axis). Cash: dotted line (right y-axis). Total's trend: gray line.

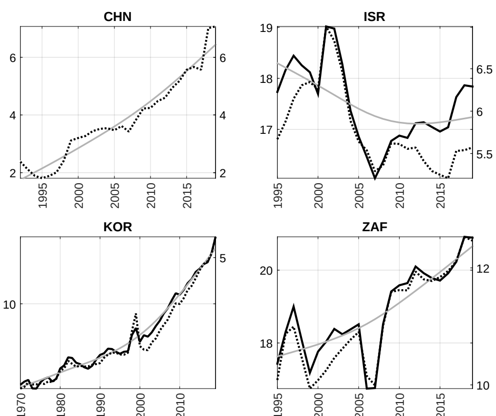
(g) Mediterranean



(h) Visegrád+Slovenia



(i) Non-Western countries



(j) Oceania

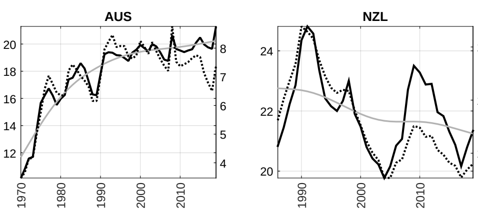


Figure 19: Social Benefits (% GDP).

Total: solid line (left y-axis). Cash: dotted line (right y-axis). Total's trend: gray line.

D.2 Social benefits versus business policies

In this section, I compare the evolution of spending on social benefits versus spending on business policies in the European Union (EU). The main finding is that there is a negative relationship over time between spending on social benefits and business policies (as a fraction of GDP). This serves as preliminary evidence to support the trade-off between social benefits and business policies that the candidate government faces in the model.

Government policies that affect businesses can be broadly divided into macroeconomic and microeconomic policies. The first category involves areas such as monetary policy, fiscal policy, and trade policy. Microeconomic policies include mainly regulatory policy and industrial policies. The type of business policies studied in the model can be interpreted as industrial policies, which comprise measures such as direct transfers to firms, low-interest loans, subsidised credit to SMEs, and R&D aid to businesses.

Constructing a single indicator that encompasses the wide variety of measures that constitute industrial policy is quite challenging. Thus, I focus on a specific aspect of industrial policy, namely state aid to industry. State aid to industry corresponds to financial transfers provided by the government to businesses, which closely aligns to the specific policy instrument studied in the model.

Data on state aid is scarce, likely due to its lack of international acceptance, such as by the World Trade Organization. Thus, the data presented covers only the countries from the EU. The unique institutional feature of the EU is that all the member states agreed to have their state aid activities monitored by the European Comission (EC). The data is obtained from the State Aid Scoreboard of the EC. I follow Stöllinger and Holzner (2017) and construct a measure for state aid to industry by adding up the spending on the following categories: commercialization, sectoral and regional development, training, internationalization, SMEs including risk capital, employment, R& D, and environmental.

Figure 20 depicts the change in spending on social benefits versus spending on industrial policies as a fraction of GDP. The x-axis corresponds to the average percentage change in spending on business policies from 2000 to 2019. The y-axis shows the same for spending on social benefits. To account for time-specific shocks that can alter the mean percentage change (e.g. the 2008 Financial Crisis), I consider the residuals after controlling for year fixed effects. The figure remains qualitatively similar when using the raw data.

Overall, countries can be categorized into two main groups based on their spending patterns. The first group of countries has been increasing spending on social benefits while simultaneously decreasing spending on industrial policies, as indicated by the red squares. On the other hand, the second group has been decreasing spending on social benefits but increasing spending on industrial policies, as indicated by the blue triangles. There are also some countries (black dots) that do not fall into either of these categories. However, in general, there is a negative relationship over time between spending on social benefits and industrial policies, as shown by the negative slope of the dotted line in the figure.

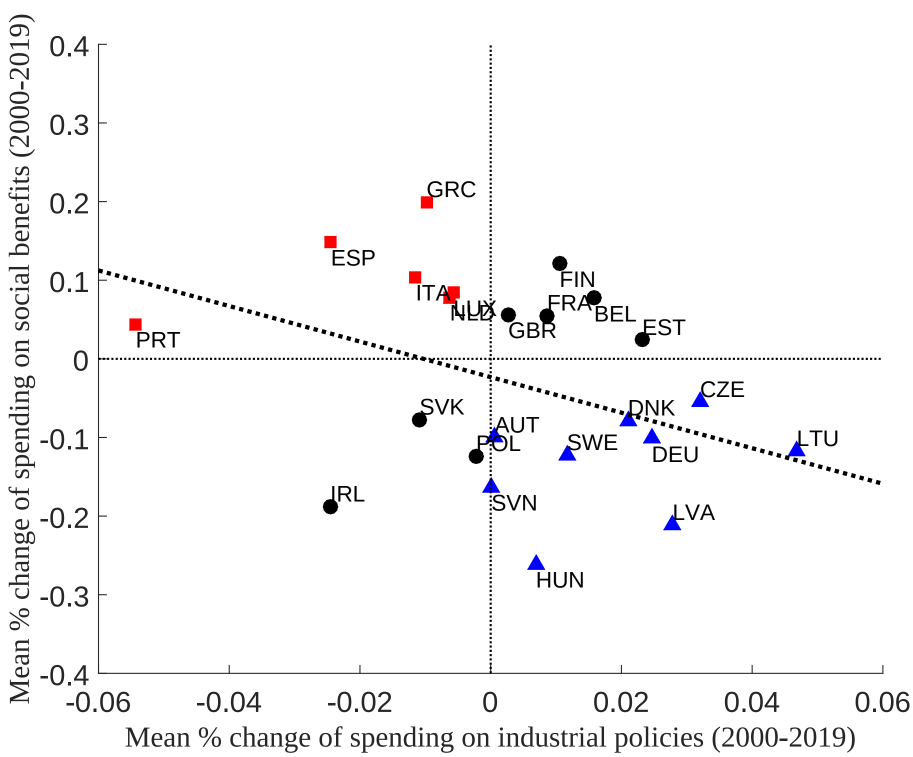


Figure 20: Spending on social benefits versus spending on industrial policies.

E Appendix: Quantitative Exercise

E.1 Input data

The quantitative exercise requires three country-level measures: spending on social benefits as share of GDP, the observed wealth distribution, and the path of aggregate productivity in the time-horizon considered.

I select countries with available data starting from 1995 or earlier, ensuring a minimum of 25 years of data. The countries considered in the simulation exercise include: Australia, Belgium, Canada, Denmark, Spain, Finland, France, Great Britain, Germany, Greece, Hungary, Ireland, Israel, Italy, South Korea, Latvia, Luxembourg, Norway, New Zealand, Poland, Portugal, Sweden, United States, and South Africa.

The starting year of each country corresponds to the earliest available year for which there is information for all variables. For most of the countries the initial year is $T_0 = 1995$, except for the United States ($T_0 = 1970$) and France ($T_0 = 1980$). To rule out the effects of the pandemic, the final year of the simulation exercise is $T = 2019$.

E.1.1 Social benefits

The data on social benefits comes from the "social benefits to households" indicator available at the OECD National Accounts and Statistics database. This source offers "reliable and internationally comparable statistics on public and (mandatory and voluntary) private social expenditure at programme level". I use the spending in social benefits as a share of GDP.

E.1.2 Wealth distribution

The calibration and simulation of the model requires country-level data on the wealth distribution. I use the World Inequality Database (WID) which provides the most extensive database on the evolution of wealth distribution across countries. The database provides detailed percentile data for the households' average net personal wealth which I use to recover the starting wealth distribution for each country, denoted as Γ_0.

E.1.3 Production function and aggregate productivity

The simulation exercise requires estimating country-level production functions and the time series for aggregate productivity. The output-production function is given by: $F(K_t, L_t) = Z_t K_t^{\alpha} L_t^{1-\alpha}$. The simulation requires estimating α and $\{Z_t\}_{T_0}^{T}$ for each country. In the baseline estimation, I use the Solow residual approach, which assumes that α is equal to the share of capital income in national income. I use the data on GDP, capital, labor, and labor income share from the Penn World Table to recover α and $\{Z_t\}_{T_0}^{T}$ for each country.

Alternatively, to account for simultaneity and selection problems, I use the method proposed by (Olley and Pakes, 1996) to estimate production functions. This procedure also allows to recover the time series for aggregate productivity across countries. The estimation requires firm-level data across countries for sales, capital, labor, and investment. The data comes from Compustat North America and Compustat Global.[46]

E.2 Calibration

E.2.1 Common parameters

The data is in annual base. For computational purposes, each year is divided into 20 periods. Thus, the length of a period is $\Delta = 0.05$.

The relative risk aversion coefficient σ is set to 2, as is common in the macroeconomic literature. Households cannot borrow, thus the minimum borrowing limit is $\underline{a} = 0$. Equation (5.7) implies that the minimum sustainable transfer rate is $\underline{b} = \frac{(\varphi r-1)I}{Y_{PE}}$, which depending on the country-specific parameters φ, r and I may be negative. To sum up, the common parameters across countries are $\Delta = 0.05, \sigma = 2$, and $\underline{a} = 0$.

[46]The countries that have enough firm-level data available to apply this estimation procedure includes Australia, Belgium, Canada, China, Germany, Denmark, Finland, France, Great Britain, Greece, Israel,Netherlands, Norway, Sweden, United States. This alternative approach gives a prediction rate of 70%.

E.2.2 Country-specific parameters

There are seven country-specific parameters that are calibrated while taking as given the observed wealth distribution of each country at the year T_0. These parameters include the interest rate (r), the discount factor (ρ), the political weight (ϕ), the fixed investment (I), the return on investment (R), the labor endowment (ℓ), and the government responsiveness parameter (ω).

E.2.3 Calibration strategy

I use eight moments to calibrate the seven country-specific parameters: $\Psi = (r, \rho, \phi, I, R, \ell, \omega)$. The traditional approach used in the macro literature is to calibrate the parameters of the model based on the moments of a stationary wealth distribution. However, in my simulation exercise the starting wealth distribution Γ_0 is taken as given from the data, so I have to opt for a different calibration strategy. In particular, I match four moments at T_0 and four moments at $T_0 + \Delta$.

The most important moment at T_0 is the share of social benefits as a function of GDP. In the model, it corresponds to the initial transfer rate: $P(\Gamma_0, \Psi)$. I also match the capital labor ratio ($K(\Gamma_0, \Psi)/L(\Gamma_0, \Psi)$), the investment to output ratio ($I \cdot e^*(\Gamma_0, \Psi)/Y_{PE}(\Gamma_0, \Psi)$), and the income Gini ($Giniy(\Gamma_0, \Psi)$).

The four moments that are matched at $T_0 + \Delta$ help to maintain the stability of the model at T_0, that is, that in the neighborhood of T_0 the economy does not experience drastic deviations from its initial state. These moments also help to identify the discount rate (ρ), and the government responsiveness parameter (ω), that do not affect the moments at T_0, but that have a large impact on the moments at $T_0 + \Delta$. The moments considered are the transfer rate ($P(\Gamma_\Delta, \Psi)$) plus three moments of the next period wealth distribution: the aggregate wealth ($\mathbb{E}[a|\Gamma_\Delta]$), the variance ($VAR[a|\Gamma_\Delta]$), and the Gini coefficient ($Gini[a|\Gamma_\Delta]$).[47]

[47] The next period wealth distribution Γ_Δ depends on the set of parameters Ψ and the initial wealth distribution Γ_0. I omit the dependence on both measures to simplify notation.

The moments considered are compactly written as:[48]

$$m(\Psi|\Gamma_0) = \begin{bmatrix} b_0 - P(\Gamma_0,\Psi) \\ K_0/L_0 - K/L(\Gamma_0,\Psi) \\ I_0/Y_0 - Inv(\Gamma_0,\Psi) \\ Giniy_0 - Giniy(\Gamma_0,\Psi) \\ b_0 - P(\Gamma_\Delta,\Psi) \\ \mathbb{E}[a|\Gamma_0] - \mathbb{E}[a|\Gamma_\Delta] \\ Var[a|\Gamma_0] - Var[a|\Gamma_\Delta] \\ Gini[a|\Gamma_0] - Gini[a|\Gamma_\Delta] \end{bmatrix}_{8x1} \tag{E.1}$$

where $Inv(\Gamma_0,\Psi) \equiv I \cdot e^*(\Gamma_0,\Psi)/Y_{PE}(\Gamma_0,\Psi)$ and $K/L(\Gamma_0,\Psi) \equiv K(\Gamma_0,\Psi)/L(\Gamma_0,\Psi)$. Thus, to find the country-specific set of parameters $\hat{\Psi}$ I solve the following problem:[49]

$$\hat{\Psi} = \arg\min_{\Psi}\{m(\Psi|\Gamma_0)'\, W\, m(\Psi|\Gamma_0)\}, \tag{E.2}$$

where W is a diagonal $8x8$ weight matrix which puts a significantly larger weight on the first moment.[50]

E.3 Simulation procedure

The simulation exercise receives as input the empirical wealth distribution, Γ_0, and the path of aggregate productivity, $\{Z_t\}_{T_0}^T$. Given $(\Gamma_0, \{Z_t\}_{T_0}^T)$, I use the consumer's decision rules in Lemma 7

[48] $b_0, K_0/L_0, I_0/Y_0$, and $Giniy_0$ correspond to the initial social benefits, capital to labor ratio, investment to output ratio, and the Gini income index from the data. $\mathbb{E}[a|\Gamma_0]$, $Var[a|\Gamma_0]$, and $Gini[a|\Gamma_0]$ are the mean, the variance, and the Gini index of the empirical wealth distribution at the starting year.

[49] I restrict the interest rate to $r \geq 0$, the political weight to $\phi \geq 1$, and the government responsiveness parameter to $\omega \in [0,1]$.

[50] Specifically, I use $W = \text{diag}(3, 0.01, 0.01, 0.01, 1, 0.5, 0.5, 0.5)$. The main objective of the calibration process is to match as closely as possible the observed transfer rate at the starting year. This is the reason why I give a significantly larger weight to the first and fifth moments. A secondary objective is the stability of the wealth distribution around T_0. Thus, I give relatively high weights to the last three moments. I explored alternative weight matrices to see whether they can have a significant impact on the generated path of social benefits. In general, changing the weight matrix can change the selected parameters. However, as long as W allows the model to match the initial transfer rate, the predicted trend of social benefits does not change.

to simulate the behavior of 1 million agents over $(T - T_0) \cdot \Delta^{-1}$ periods.[51] This is equivalent to iterating forward the KF equation (3.15). Occupational decisions are governed by equation (3.12), while the equilibrium transfer rate is determined by (5.7) and (7.1).

The first step in the simulation algorithm is to obtain a draw from the empirical wealth distribution in the starting year, Γ_0. I recover Γ_0 based on country-level percentile data for households' average net personal wealth (taken from WID). Then, I adjust a Pareto distribution to Γ_0 and proceed to obtain a draw of 1 million agents.[52] After obtaining a draw of the initial distribution, the simulation algorithm amounts the following iterative procedure:

1. Given Γ_t and Z_t, obtain the optimal fraction of entrepreneurs, e_t^*. Use equation (7.1) to obtain e_t.

2. Given Γ_t and e_t, obtain the occupational condition by solving: $a_t^o = \Gamma_t^{-1}(1 - e_t)$.

3. Given a_t^o, use equation (5.7) to obtain the equilibrium transfer rate b_t.

4. Use the policy rules from Lemma 7 and a_t^o to simulate the behavior of the 1 million agents. This step gives rise to the next period wealth distribution $\Gamma_{t+\Delta}$.

E.4 Quantitative results

Figure 21 compares the observed trend of social benefits (solid line, left y-axis) with the trend predicted by the model (dotted line, right axis) for the 24 evaluated countries.[53] The "X" symbol to the right of some country names indicates that the model fails to predict the trend observed in the data. The countries that the model cannot explain include: Belgium, Denmark, Finland,

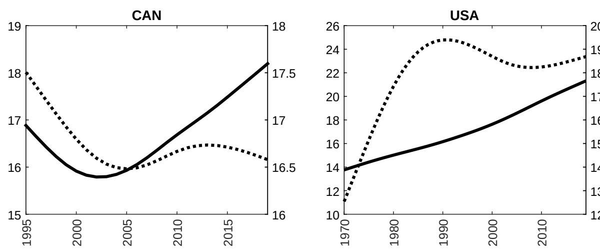
[51] For the majority of the countries, the length of the simulation exercise is 500 periods: $(2019 - 1995 + 1) \cdot 20$. For the United States and France, I simulate the model for 1000 and 800 periods, respectively.

[52] The minimum borrowing limit is $\underline{a} = 0$, so I apply this restriction to Γ_0. Some countries may have negative assets in the first percentiles. Thus, the minimum wealth level in those countries is zero, which makes the adjustment to a Pareto distribution infeasible. To solve this problem, I shift the wealth distribution to the right by a fixed amount, and then, I proceed to adjust a Pareto distribution. The amount by which Γ_0 is shifted to right affects the parameters of the adjusted Pareto distribution. I choose this "shift" to maintain the mean, the variance, and Gini index as close as possible to the values of the initial wealth distribution, Γ_0.

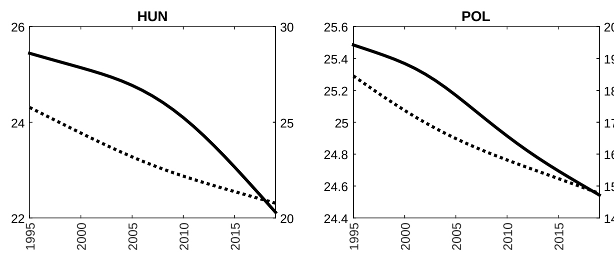
[53] In some cases, the predicted trend's magnitude might differ from the actual data. To facilitate comparison, the data and the model results are presented on separate axes.

France, Ireland, and Portugal. Overall, the model can predict the observed trend of social benefits for 18 out 24 calibrated countries.

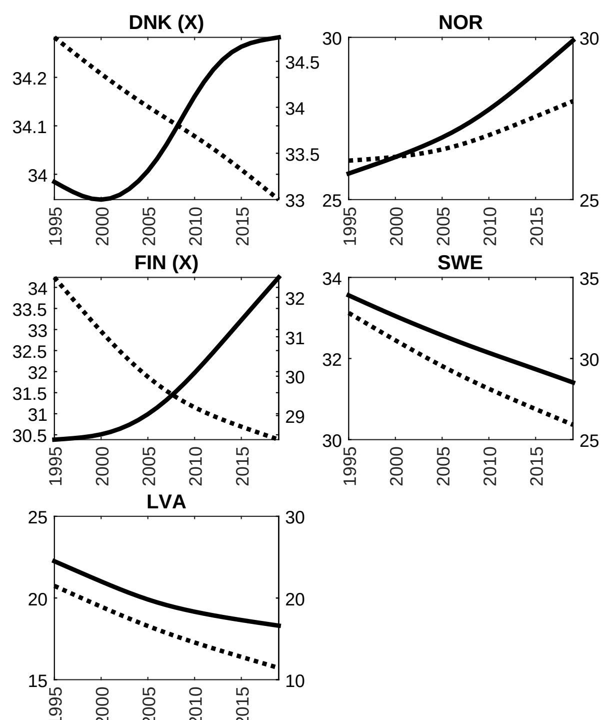
(a) North America



(b) Eastern Europe



(c) Northern Europe



(d) Western Europe

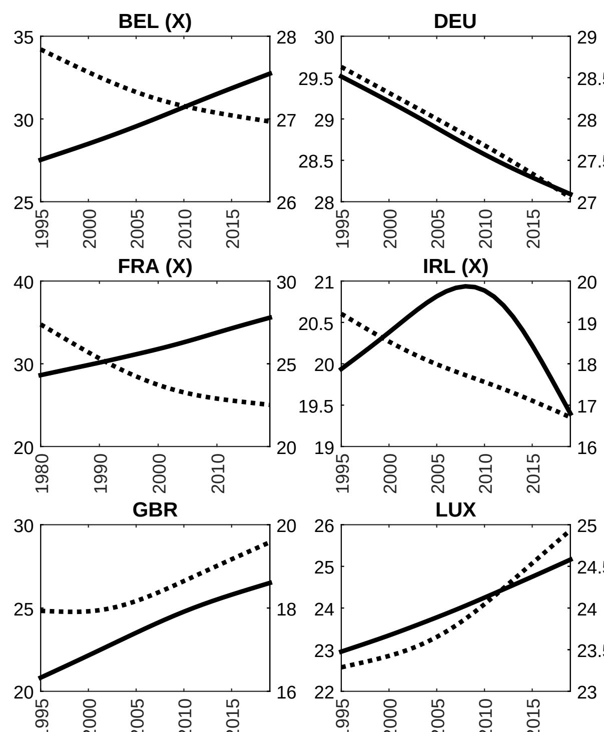
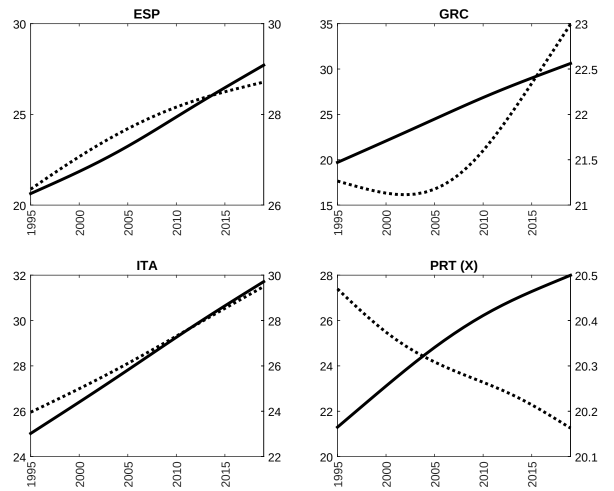


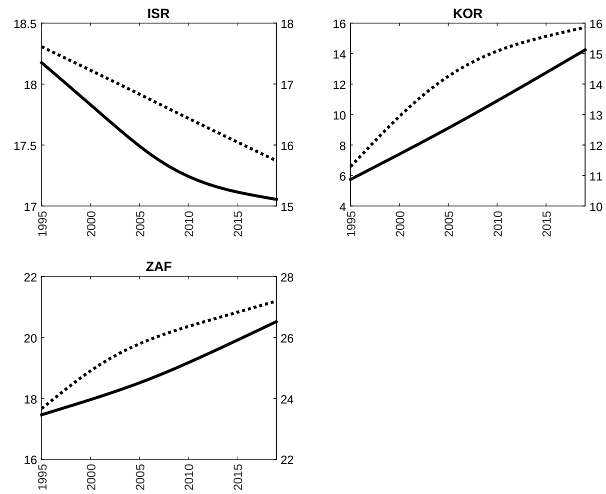
Figure 21: Social Benefits (% GDP): observed versus predicted trends

Data: solid line (left y-axis). Model: dotted line (right y-axis).

(e) Southern Europe



(f) Non-Western countries



(g) Oceania

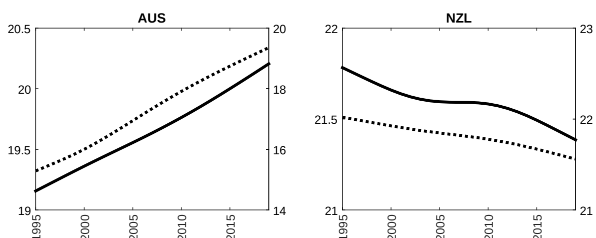


Figure 21: Social Benefits (% GDP): observed versus predicted trends

Data: solid line (left y-axis). Model: dotted line (right y-axis).

E.5 Counterfactual analysis

The objective of this section is to evaluate the role of changes in the government political orientation (ϕ) in the evolution of the size of the welfare state (b). I consider three countries: Canada, the United States, and Sweden. For each of these countries, I solve for the path of political weights that matches the observed path of social benefits. Then, I simulate a highly pro-business (high ϕ) and a highly pro-worker scenario (low ϕ). The main finding is that the trend of social benefits would not have changed significantly in either of the three countries and under both scenarios.

The counterfactual exercise confirms the key role of the wealth distribution in explaining the dynamics of the welfare state.

E.5.1 Calibration: Searching for political weights

In the baseline model, the political weight ϕ is fixed over time within countries. In the counterfactual analysis, I allow the political weight to follow an exogenous path in each country, $\{\phi_t\}_{T_0}^{T}$, to capture changes in the political orientation of the government.

The first step is to find the sequence $\{\phi_t\}_{T_0}^{T}$ that matches the path of social benefits in the data, $\{b_t\}_{T_0}^{T}$. The rest of the parameters are fixed for each country. Their values are obtained from the calibration approach used for the main quantitative exercise in Section 7.

For computational purposes, I consider a discrete time model where the length of a period is $\Delta = 0.05$. Thus, each year is divided into 20 periods. The data is annual base, so I allow the political weight to change every one year, i.e. I search for $T - T_0 + 1$ values for each country.

The calibration procedure relies on the political equilibrium (PE) condition and the KF equation (3.15) evaluated at the different values of ϕ. The PE condition is the implicit function defined by (5.4) and (7.1) that provides a mapping from the wealth distribution Γ_t to the equilibrium transfer rate, denoted in this section by: $P(\Gamma_t; \phi_t)$.

Given the initial observed wealth distribution, Γ_0, I obtain a draw of 1 million agents. After obtaining a draw for the initial distribution, the calibration algorithm amounts the following iterative procedure:

1. Given Γ_t and Z_t, solve the following minimization problem: $\phi_t = \arg\min_{\phi \geq 1}\{|b_t - P(\Gamma_t; \phi)|\}$.

2. Given the equilibrium transfer rate $P(\Gamma_t; \phi_t)$, use the policy rules from Lemma 1 to simulate the behavior of the 1 million agents. This step gives the next period wealth distribution Γ_{t+1} as a function of the current wealth distribution Γ_t and the political weight ϕ_t.

E.5.2 The evolution of the political weight

For space considerations, in Figure 22, I present the estimated political weights only for the US. The vertical gray lines separate the different administrations over time. The blue dashed line represents Republican governments, while the red dotted line corresponds to Democrats.

Panel a) shows the estimated path for the political weight. There is a clear decreasing trend since 1970, that is, governments have become more pro-worker in the last fifty years. Panel b) depicts the average change of the political weight by administration. In general, the larger increases of ϕ coincide with Republican governments, consistent with a more pro-business partisan nature. On the other hand, the largest decreases of ϕ happen during Democrat administrations indicating a more pro-worker stance.

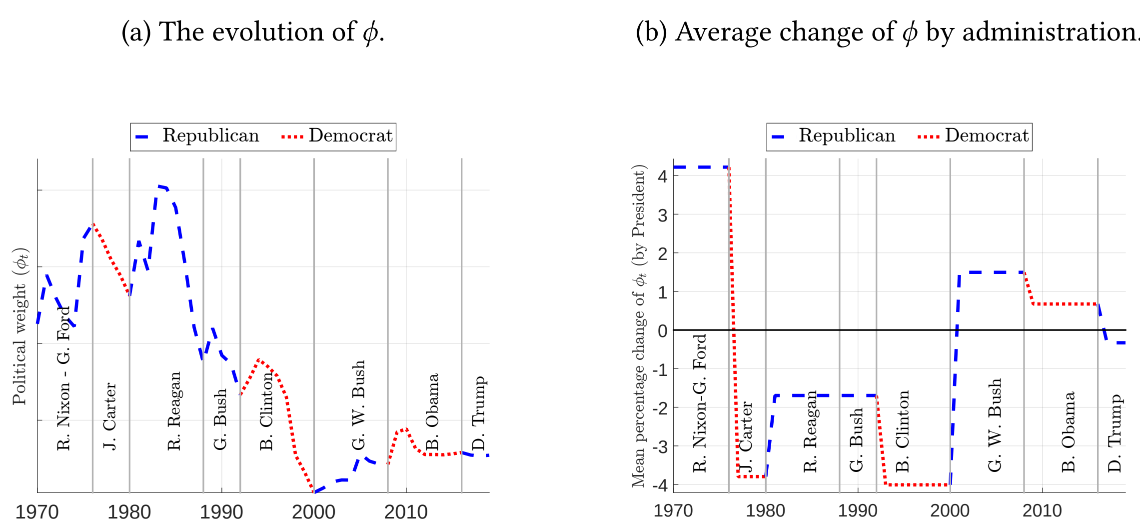


Figure 22: USA: The estimated path of political weights (ϕ).

E.5.3 Simulation: A pro-worker versus a pro-business scenario

Figure 23 presents the evolution of social benefits for the US under both counterfactual scenarios. The solid line is the data. The red dotted line represents the pro-worker scenario which is constructed by multiplying the estimated path of political weights by the largest percentage decrease in Figure 22. The blue dashed line is the pro-business scenario which is constructed by multiplying ϕ_t by the largest percentage increase across administrations. The gray lines are the trends.

Under both "extreme" scenarios, the trend of social benefits would have remained positive since 1990. For Canada and Sweden, the trends observed in the data do not change under both scenarios. Thus, changes in government administrations play a limited role in shaping the evolution of the welfare state.

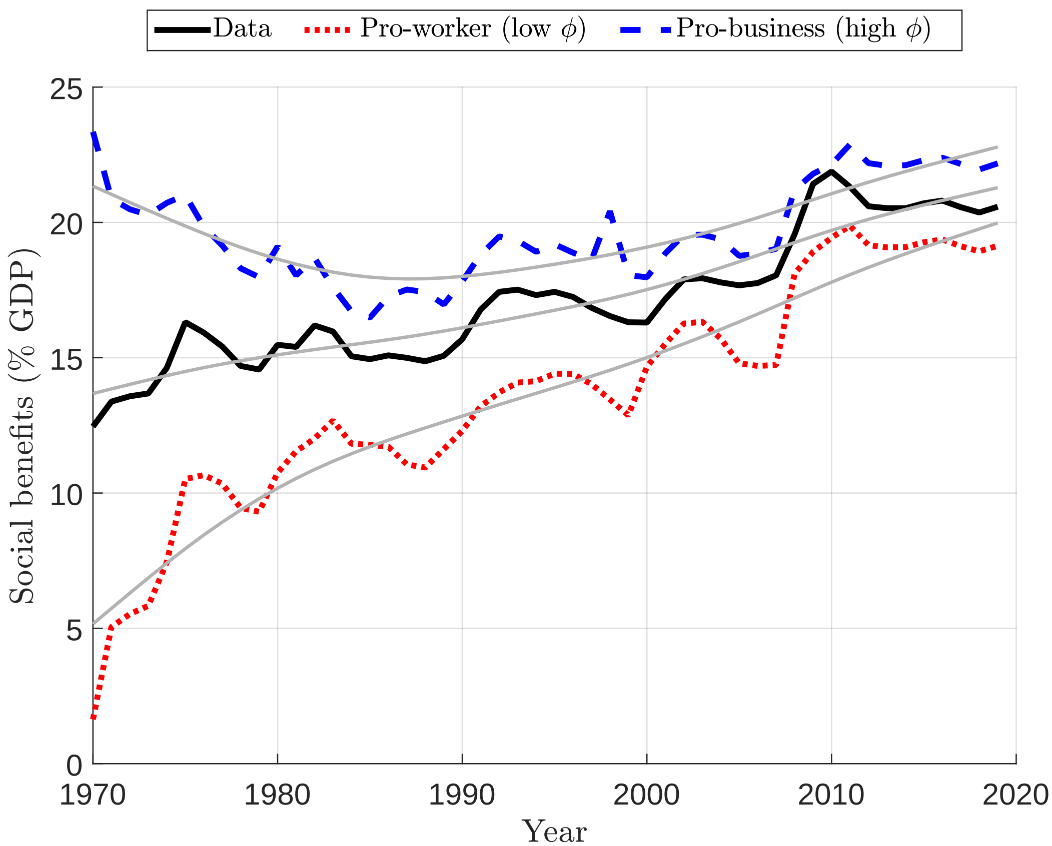


Figure 23: USA: Counterfactual scenarios.

F Appendix: Additional Figures

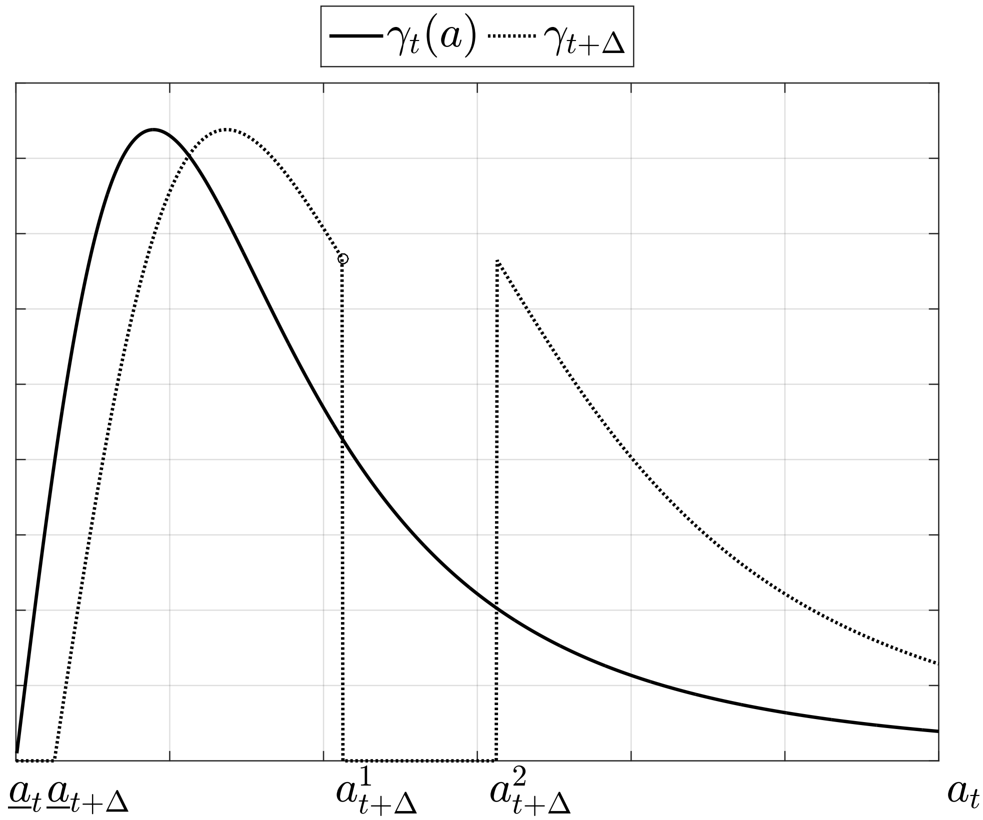


Figure 24: A right shift of the wealth distribution function, $\gamma_t(a)$.

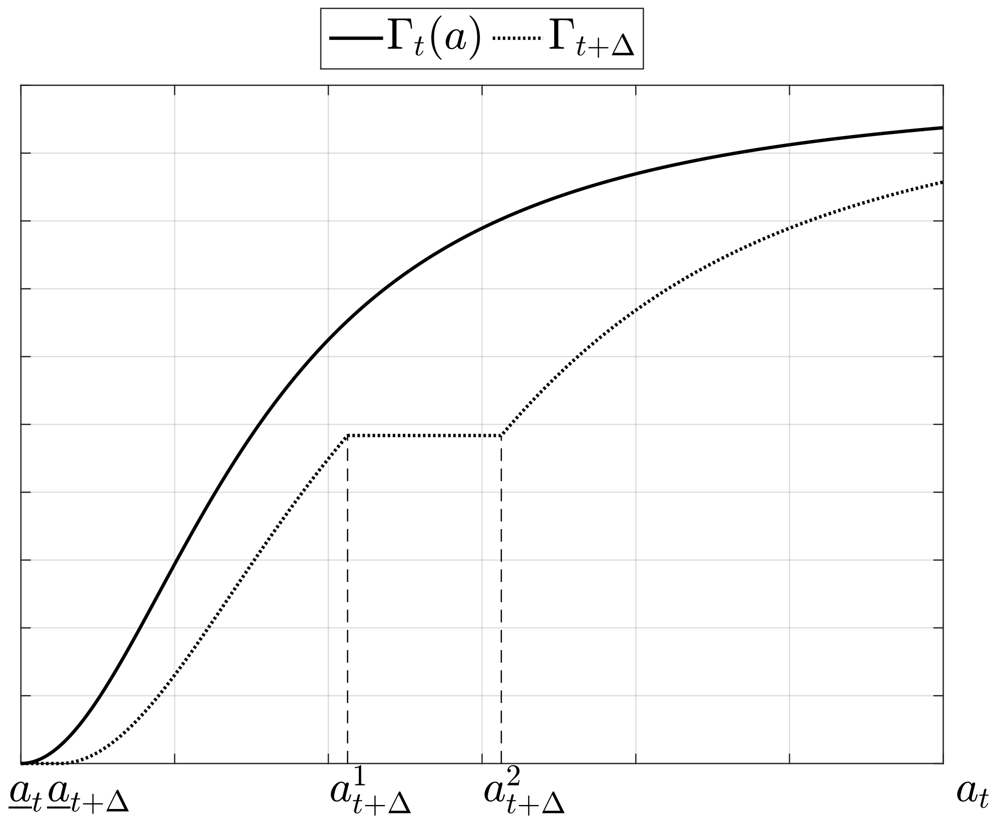


Figure 25: A right shift of the cumulative wealth distribution function, $\Gamma_t(a)$.

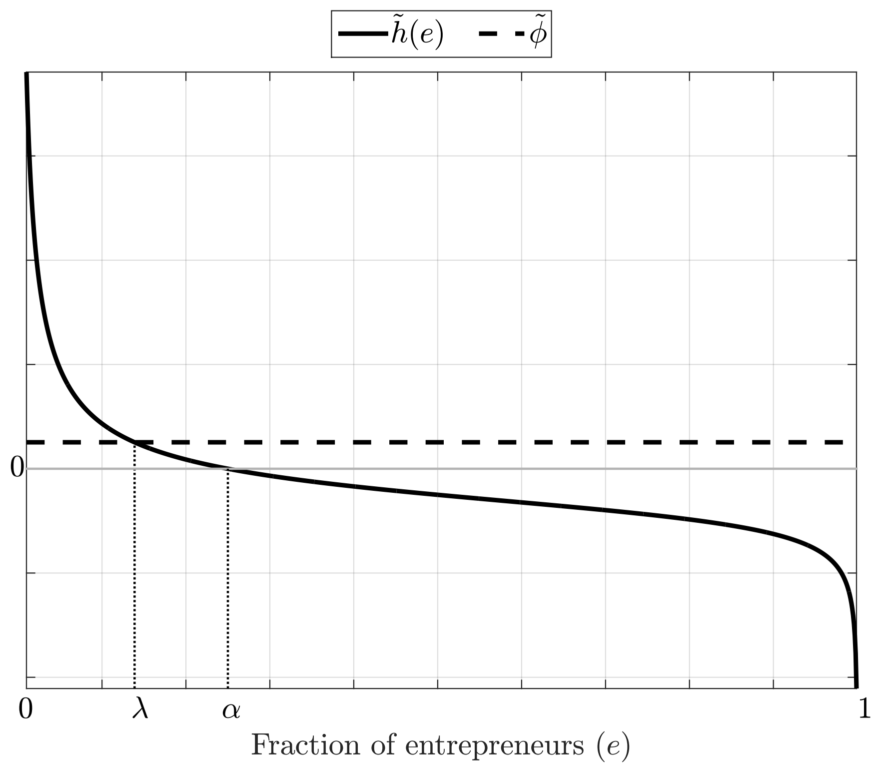
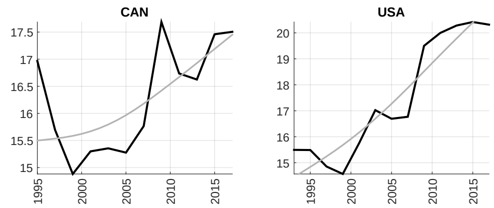
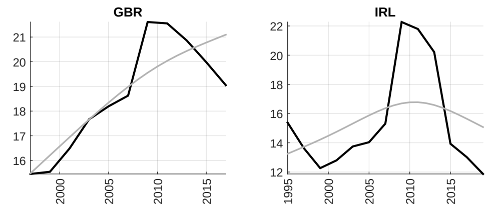


Figure 26: The political equilibrium condition.

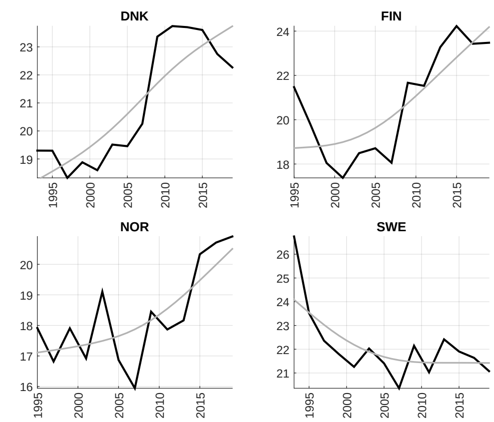
(a) North America



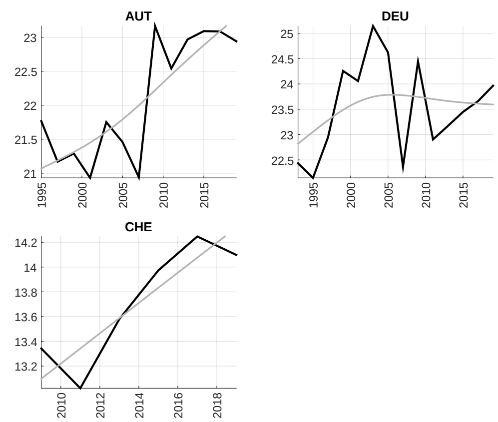
(b) British Isles



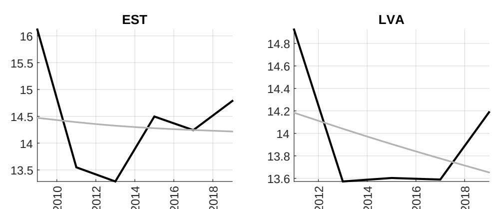
(c) Scandinavia



(d) DACH



(e) Baltic States



(f) Benelux

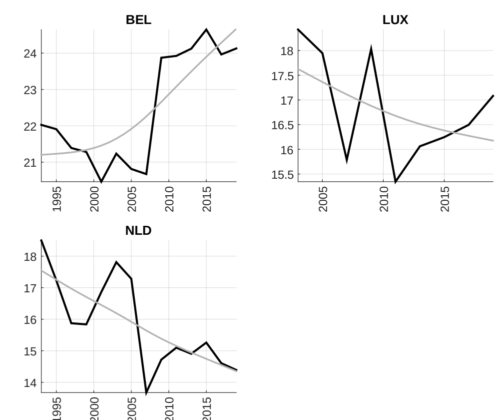
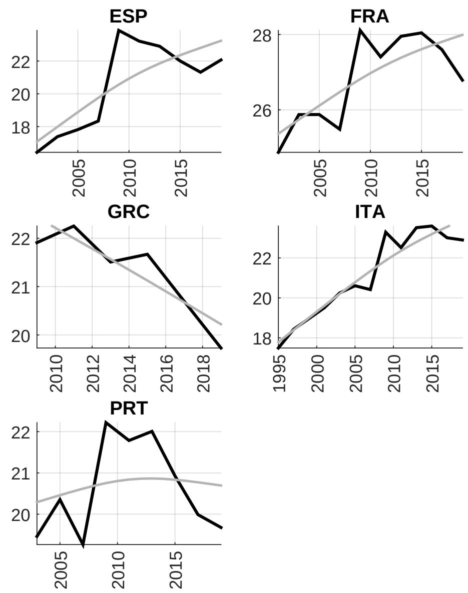


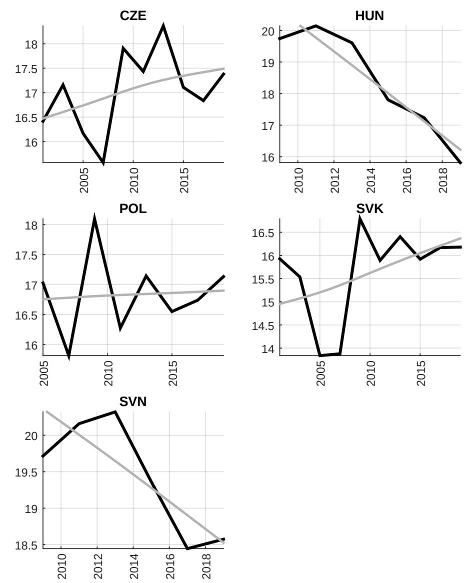
Figure 27: Net public social spending (% GDP).

Total: solid line (left y-axis). Total's trend: gray line.

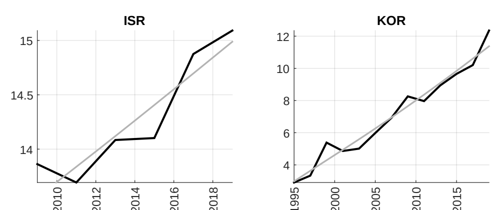
(g) Mediterranean



(h) Visegrád+Slovenia



(i) Non-Western countries



(j) Oceania

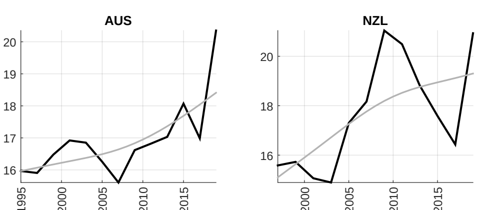


Figure 27: Net public social spending (% GDP).

Total: solid line (left y-axis). Total’s trend: gray line.