European University Institute

DEPARTMENT OF ECONOMICS

Essays on Income Inequality, Political Inequality and Income Redistribution in the U.S.

Pavel Brendler

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 2 June, 2016

European University Institute
Department of Economics

Essays on Income Inequality, Political Inequality and Income Redistribution in the U.S.

Pavel Brendler

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Andrea Mattozzi, EUI, Supervisor
Prof. Árpád Ábrahám, EUI
Prof. Dean Corbae, Wisconsin School of Business
Prof. Hamish Low, University of Cambridge

Researcher declaration to accompany the submission of written work

I, Pavel Brendler, certify that I am the author of the work "Essays on income inequality, political inequality and income redistribution in the U.S." I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297).

The copyright of this work rests with its author. [quotation from it is permitted, provided that full acknowledgement is made.] This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

Statement of inclusion of previous work (if applicable):

I confirm that chapter "Voter mobilization and electoral competition" was jointly co-authored with Ilya Archakov and I contributed 50% of the work.

Signature and Date:

Pavel Brendler
2nd May 2016

Abstract

The **first** chapter, entitled "Life-time earnings inequality and income redistribution through social security in the U.S.", is devoted to social security, which is the public pension system in the U.S. In 2015, the Social Security Advisory Board proposed to the Congress to change the parameters of the pension benefit formula, which have been constant since 1977. This change implies a fall in statutory replacement rates for individuals with high lifetime earnings. I construct a model, which accounts for a significant portion of the proposed change. Counterfactual experiments suggest that increased uncertainty in the labor productivity process and the upward shift in the college premium explain most of the change in the parameters of the pension benefit formula.

The **second** chapter, "Income inequality and political inequality in the U.S.", is devoted to the median voter theorem, which is one of the most celebrated results in the public choice theory. Existing structural models predict too high income redistribution for the U.S. economy if the tax rate is chosen by the median voter. One potential explanation is that the political process in the U.S. is biased towards richer agents. In this case, the decisive agent is richer than the standard median voter and therefore prefers lower redistribution consistent with the data. I introduce wealth-weighted majoritarian voting over progressive income taxation into a heterogeneous agent model with idiosyncratic risk. I show that the model can significantly better explain the dynamics of income redistribution in the U.S. since 1980s than a model with a standard median voter.

In the **third** chapter of my dissertation, entitled "Voter mobilization and electoral competition", which is a joint work with Ilya Archakov, we analyze the impact of voter abstention on electoral competition in the U.S. We present a novel game theoretic approach to study the competition between two candidates for a seat in a legislature, when candidates can spend money both on advertising to gain a larger share of potential supporters and on voter mobilization to bring the supporters to the voting poll. We show that the results of our model are consistent with the campaign expenditure data by the Federal Election Commission for the 2010 and 2012 House of Representative election cycles.

Key words: electoral competition, income redistribution, median voter, political inequality, public pension, social security, voter abstention, voter mobilization.

JEL: D72, E62, H24, H31, H55.

Acknowledgements

Introduction

This dissertation builds around three important economic topics. Below I discuss each of them in turn.

The **first** chapter of the dissertation, entitled "Life-time earnings inequality and income redistribution through social security in the U.S." is devoted to social security, which is the public pension system in the U.S. One of the key features of public pension systems in many countries is that individual's pension benefit depends on individual's average lifetime earnings. This relationship is governed by a statutory pension benefit formula. The pension benefit formula is important because – through its impact on (per period) pension benefits – it affects the amount of income redistribution over individual's lifetime. It is, therefore, important to understand how the governments set the redistributive design of the public pension system.

I analyze the social security system in the U.S. in the period 1977-2015. In 1977, the U.S. Congress introduced the Social Security Amendments, which fixed the parameters of the pension benefit formula. These parameters have been kept unchanged since then. In 2015, however, the Social Security Advisory Board, which advises the Congress on the problems at the Social Security Administration, has proposed to adjust the parameters of the pension benefit formula and weaken the mapping from lifetime earnings to pension benefits for households with high and medium average lifetime earnings.

I use the data by the Social Security Administration to argue that the Congress adopted the Amendments of 1977 at a time, when the inequality in average lifetime earnings among workers was relatively low. I show that following an increase in cross-sectional earnings inequality since 1970s, the inequality in average lifetime earnings across workers has increased significantly. If there is substantial heterogeneity in lifetime earnings among retired households, the redistributive design of the pension system is likely to have large welfare effects for households. To analyze these effects, I build a model, in which agents are heterogeneous with respect to their lifetime earnings and in which a social planer decides upon the parameters of the pension benefit formula. The model predicts a fall in statutory replacement rates for retirees with medium to high average lifetime earnings and an increase for retirees with low average lifetime earnings. At the same time, the model predicts an increase in the overall size of the pension program because the social

security tax rate rises. I use the model to translate these changes in statutory replacement rates into the implied changes in ex-post lifetime replacement rates and compare them for households born before and after the reform in 2015. The model predicts an increase in lifetime replacement rates of 2.4 percentage points for the households from the bottom quintile and a drop of 1.9 percentage points for the households from the top quintile of the lifetime earnings distribution. I find that the increase in cross-sectional earnings inequality drives most of the change in lifetime income redistribution.

In the **second** chapter, "Income inequality and political inequality in the U.S.", I quantify the degree of political inequality of the American society. I extend a standard heterogeneous agents model with incomplete financial markets by introducing *wealth-weighted* voting over income redistribution. Political inequality is modeled as a simple functional relationship between private wealth holdings and agent's voting share. I set up a benchmark model economy, which matches the observed levels of earnings inequality in the U.S. in 1979. In particular, the wage process is calibrated to match relatively low dispersion of wages and relatively low mobility across wage groups in the data at that time. I use the model to identify the degree of political inequality, which is consistent with the level of income redistribution observed in late 1970s in the U.S. This allows me to uniquely pin down the state, in which the *weighted* median voter is located in the model in 1979. Then I conduct the following experiment. I reparameterize the wage process to match the increased cross-sectional earnings inequality in 1996. Afterwards, I let agents vote in 1979 once-and-for-all on the degree of the income tax progressivity. This experiment imitates the adoption of the Economic and Recovery Tax Act of 1981, which significantly reduced the progressiveness of the statutory income tax schedule. Before agents vote, they anticipate the occurred changes in the wage process. I simulate the counterfactual time series of income redistribution since 1979 under the standard and the weighted median voters and compare them with the data.

I find a very high degree of political inequality in 1979. Whereas the standard median voter is located in the 2nd quintile of the equilibrium market income and earnings distributions of 1979, the weighted median voter is in the top decile of the market income distribution and 8th decile of the earnings distribution. The model with the identified weighted median voter is able to account for roughly 36 percent of the decline in income tax progressivity due to the Economic and Recovery Tax Act of 1981. Moreover, I show that in response to the increased earnings inequality a model, in which the standard median voter is decisive, predicts twice as much redistribution at the bottom and more two and a half times more redistribution at the top than in the data. On the contrary, a model, in which the weighted median voter is decisive, generates the redistributive effects of taxes and transfers, whose magnitudes are in line with the data. Finally, in order to understand the differences between the voting outcomes under the standard and the weighted median voters, I quantify the effect of rising dispersion of wages and rising mobility across wage

groups since 1980 on the most preferred tax rate of each type of the voters. I show that agent's expectations about the future wage inequality are crucial in understanding the differences between the two equilibrium outcomes.

The **final** chapter, "Voter mobilization and electoral competition", which is a joint work together with Ilya Archakov, explores the effect of voter abstention on electoral competition. We suggest a simple model, in which candidates compete by running costly campaigns in order to increase their probabilities of winning the elections. Being agnostic about the choices of policy platforms, we focus on the campaign spending strategies of the candidates. There are two groups of voters in a district. The voters of the first type are partisan (core) supporters of the candidates. Such voters are already confirmed in their decision for which candidate to vote in case they participate in voting. Voters of the second type are independent (swing). Such voters may vote for either candidate depending on electoral campaign impact and stochastic preference fluctuations. The voters of both types, however, may either participate in voting or abstain. We assume that the participation decision is totally unrelated to the voters' political preferences and depends on some idiosyncratic random factors. During the electoral race candidates may affect swing voters' sympathies through advertizing and voter turnout through mobilization. We use this framework to illustrate several qualitative relations between equilibrium campaign spending and different characteristics of a given electoral district and given electoral race in general. A distinctive feature of our framework is that it allows us to analyze advertising and mobilization spending of candidates as parts of a single electoral campaign.

Our model reconciles persuasion and mobilization in a unified framework. Candidates act in two consecutive phases of the electoral race. In the first stage, they try to persuade independent citizens to vote for them by means of costly advertizing activities. In the second phase, candidates mobilize their core supporters after receiving a coarse signal about the state of relative popularity which has been realized after the advertizing stage. Since both advertizing and mobilization activities are costly, candidates also exert fundraising efforts which are the source of disutility in their objectives.

From the analysis of our novel modeling framework we derive a set of qualitative results. One of them establishes the link between equilibrium spending and the degree of ex-ante electoral advantage of one of the candidates, which can be interpreted as, for example, an incumbency advantage. Both advertizing and mobilization expenditures decrease as far as an ex-ante disparity between candidates rises (expected competition is low). We test the implications of our model using the administrative data on campaign expenditures by the Federal Election Commission for the 2010 and 2012 House of Representative election cycles. Using the Tobit regression model, we confirm that both advertizing and mobilization expenditures of candidates increase with expected competitiveness of the race, where we control for the expected competitiveness using the *2010 and 2012 House Race Ratings* by *The Cook Political Report.* We also confirm that the

expected turnout, which we proxy by voter turnout rates in the previous election in the same congressional district, has a positive effect on advertizing expenditures, while we don't find any significant effect of past turnout rates on mobilization expenditures.

Contents

Chapter 1

Lifetime earnings inequality and income redistribution through social security

1.1 Motivation

One of the key features of public pension systems in many countries is that individual's pension benefit depends on individual's average lifetime earnings. This relationship is governed by a statutory pension benefit formula. The pension benefit formula is important for at least three reasons. Firstly, it determines the degree of income redistribution within retired households in a given time period. If social security taxes are proportional and pension benefits are weakly tied to earnings, then the system redistributes incomes from households with high earnings to households with low earnings. Secondly, through its impact on (per period) pension benefits, the pension benefit formula affects the amount of income redistribution over individual's lifetime. Workers pay social security contributions on their earnings and receive earnings-dependent pension benefits in retirement. The amount of income redistribution over individual's lifetime depends on the difference between the contributions paid and the pension benefits received. Finally, governments can use the pension benefit formula to tackle the solvency issue of their public pension systems. By weakening the relationship between lifetime earnings and pension benefits, the government may target incomes at disadvantaged groups of retirees, without necessarily having to raise the overall size of the pension system.

These reasons suggest that the pension benefit formula might have important economic and welfare consequences. It is, therefore, important to understand how the governments set the redistributive design of the public pension system. In this paper, I look at the social security system in the U.S. in the period 1977-2015. In 1977, the U.S. Congress introduced the Social Security Amendments, which fixed the parameters of the pension benefit formula. These parameters have been kept unchanged since then. In 2015, however, the Social Security Advisory Board, who advises the Congress on the problems

1

at the Social Security Administration (SSA), has proposed to adjust the parameters of the pension benefit formula and weaken the mapping from lifetime earnings to pension benefits for households with high and medium average lifetime earnings.[1]

I use the data by the SSA to argue that the Congress adopted the Amendments of 1977 at a time, when the inequality in average lifetime earnings among workers was relatively low. I show that following an increase in cross-sectional earnings inequality since 1970s, the inequality in average lifetime earnings across workers has increased significantly. If there is substantial heterogeneity in lifetime earnings among retired households, the redistributive design of the pension system is likely to have large welfare effects for households. To analyze these effects, I build a model, in which agents are heterogeneous with respect to their lifetime earnings and in which a social planner decides upon the parameters of the pension benefit formula. I use this model to answer three questions: 1) What parameters of the pension benefit formula would the Congress choose if it had to decide on them in 2015? 2) By how much would lifetime income redistribution change across different types of households in response to this change in the statutory pension benefit formula? 3) How much of the change in lifetime income redistribution can be accounted for solely by the increase in cross-sectional earnings inequality?

To be more precise, I extend a stochastic overlapping generations model à la Huggett (1996) by introducing a utilitarian social planner, who decides once-and-for-all on the pension benefit schedule, which I model as:

pension of an individual $i = \alpha_1 \times$ (average lifetime earnings of an individual $i)^{\alpha_2}$.

I show that this representation fits fairly well the statutory pension benefit formula. For $\alpha_2 < 1$, the replacement rate, i.e. the ratio of the pension benefit to the average lifetime earnings, decreases with average lifetime earnings; if $\alpha_2 > 1$, the opposite is true. A special case with a lump-sum pension benefit emerges if $\alpha_2 = 0$. The variable α_1 shifts up and down the replacement rate curve. In the model, the government chooses α_1 and α_2 to maximize the welfare of all generations alive. Given these two variables, the social security tax rate adjusts in each period to balance the government budget.[2] The government chooses the policy under the rational expectation about the effects of these policies on future equilibrium outcome and the welfare of each agent during her entire lifetime.

This representation of the pension benefit formula allows me to study important

---

[1] See http://www.ssa.gov/retirementpolicy/projections/worker-benefit.html.

[2] The assumption of the balanced budget is less restrictive than it might first seem. The budget of the social security Trust Fund was roughly balanced until 1983 (the period, to which I calibrate the benchmark model). During 1984-2008 (the period, during which I simulate the benchmark model), the Trust Fund was running moderate surpluses. Only since 2009, the program has been recording steadily growing deficits. I discuss below how the assumption of the balanced budget might change the main results.

and non-trivial trade-offs faced by the U.S. government. The government can change the overall size of the pension system by changing α_1. At the same time, for a given size of the system, the government can control the distribution of pension benefits within retired cohorts through α_2. An increase in α_1 directly raises the total size of the pension fund and for a given α_2 increases the amount of redistribution. By choosing α_2, the government can direct incomes towards the agents with lower lifetime earnings because they have higher marginal utility of consumption. Furthermore, in the absence of formal or informal private insurance markets against idiosyncratic uncertainty, the public pension system provides a partial substitute for these missing markets and therefore may lead to less volatile household consumption over time. Combining the two policies α_1 and α_2 together, the government can achieve a more equal distribution of income, and therefore, wealth, consumption and welfare. But redistribution and insurance go hand-in-hand with distortions. A larger size of the system has to be financed through higher social security taxes, which reduce agent's individual and aggregate saving. Similarly, for a given size of the pension fund, targeting the pension benefit at the groups with higher or lower average lifetime earning distorts the savings of these groups because they anticipate a higher future transfer in retirement. Due to the distortionary effects, the pension system also affects prices of labor and capital and the government lump-sum transfers through the general equilibrium.

I calibrate the model to the U.S. data in 1977. Given the parameters of the social security system, the model matches the inequality in lifetime earnings in the data at that time. Then I endogenize the parameters of the social security system and assume that the Pareto weight is a Cobb-Douglas function of agent's average lifetime earnings and age. I calibrate the two parameters of the Pareto weight function, so that the government in the model chooses (α_1, α_2) consistent with the Social Security Amendments of 1977.

Afterwards, I simulate the model during 1977-2015, accounting for the major economic and demographic changes in the U.S. economy, which took place during this period. The first change refers to the drastic rise in cross-sectional earnings inequality. I update the age-specific college premium and adjust the dispersion of the temporary and the persistent components of the idiosyncratic labor productivity process, consistent with the data. The second change refers to falling birth rates, which resulted in a higher old-age dependency ratio (i.e. ratio of retired to the working age population). The third change is a sharp reduction in the progressiveness of the income tax schedule during 80s, mostly due to adoption of the Economic and Recovery Tax Act of 1981. Finally, I account for the fact that the earnings cap for the hospital insurance was eliminated in 1994. Along the transition, the model matches increasing inequality in average lifetime earnings in the data. Then I ask: If the Congress applies the same welfare function as in 1977, what optimal design of the social security system would it choose now, given that the lifetime earnings inequality has increased?

There are several important findings. First, with equal Pareto weights, the government in 1977 opts for a strongly redistributive pension system. To be consistent with the Amendments of 1977, the government must have put more weight on earnings-rich households in their 40s. Second, I find that those who are born in 2015 shortly before the change of the pension system receive negative net transfers (pension benefits net of paid-in social security contributions) as compared to the newborns in 1977: -9% for the agents from the first quintile and -12% for the agents from the top quintile of the lifetime earnings distribution. I find that the optimal design of the pension system in 2015 implies a fall in statutory replacement rates for retirees with medium to high average lifetime earnings and an increase for retirees with low average lifetime earnings. At the same time, the model predicts an increase in the overall size of the pension program because the social security tax rate rises.

Then I conduct several counterfactual experiments to study the effect of each change in the economic environment along the transition on the equilibrium pension benefit formula in 2015. I find that the increased uncertainty in the labor productivity process (i.e. increased variance of the transitory and permanent shocks) and the upward shift in the college premium explain most of the change in the parameters of the pension benefit formula.

The paper builds upon the macroeconomic literature, which develops politico-economic models to rationalize the size of welfare programs. Rios-Rull and Krusell (1999), Hassler et al. (2007), Corbae et al. (2009) apply the median voter theorem or a social welfare function as a social choice mechanism to study the interaction between income inequality and income redistribution. These studies focus on personal income taxes and government spending on welfare programs. Much less attention has been devoted to government spending on public pension benefits. One notable exception is Conesa and Krueger (1999), who build upon the general equilibrium overlapping generations model with production, incomplete financial markets and idiosyncratic labor productivity risk by Huggett (1996). The authors analyze the quantitative role of idiosyncratic risk, when agents vote once-and-for-all on a gradual elimination of social security in a simple absolute majority election. Gonzalez-Eiras and Niepelt (2008) ask the question, by how much the size of the pay-as-you-go pension system in the U.S. will rise in response to falling fertility rates. One finding is that in response to the projected demographic transition, social security tax rates will not have to increase by as much as projected by the SSA because labor supply might be rising along the transition.

Below I proceed as follows. The key data facts regarding the redistribution of pension benefits and the evolution of lifetime earnings inequality are discussed in section 1.2. I set up the model in section 1.3. Section 1.4 calibrates the model, while section 1.5 describes the quantitative exercise. The results are shown in section 1.6. Section 1.7 discusses challenging but important tasks for future research.
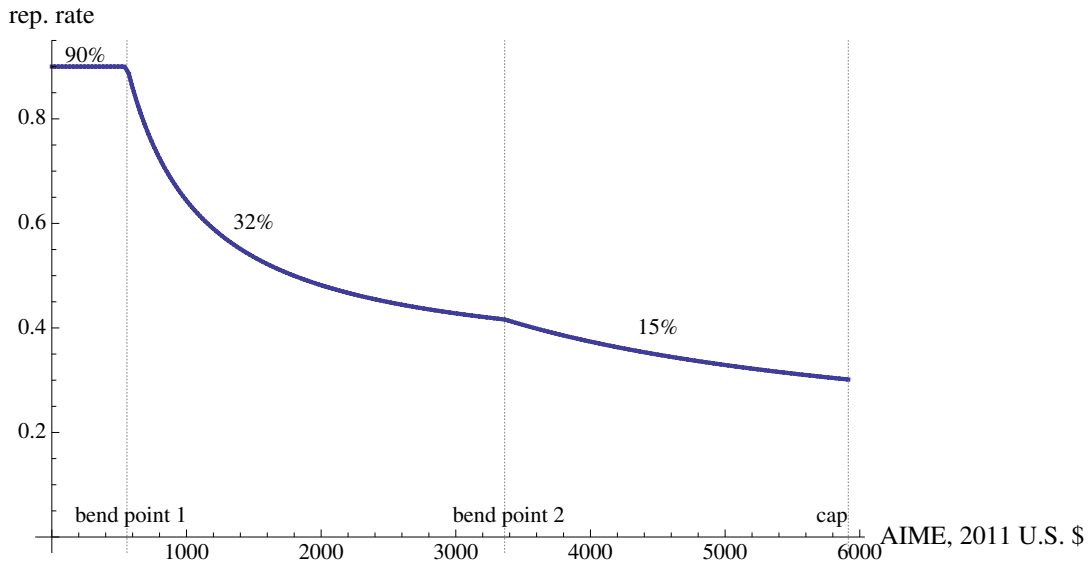
Figure 1.1: Statutory replacement rate for workers retiring in 2012

1.2 Data

1.2.1 Data fact 1: Pension benefit formula

The Social Security Amendments of 1977 fixed a specific formula and its parameters, which the SSA applies to compute the pension benefit. The key variable of this formula are the average indexed monthly earnings (AIME). Essentially, these are the average monthly earnings (adjusted for inflation and growth in wages) over individual's 35 highest years of working career.[3] I will refer to the AIME as the average lifetime earnings. Only the earnings below a certain threshold flow into the computation of the average lifetime earnings. This threshold determines the *annual maximum taxable earnings*, which I refer to as the *cap* below. The pension benefit formula then maps the average lifetime earnings into the pension benefit. More specifically, the formula multiplies a 90, 32, or 15 percent factor by the portion of worker's average lifetime earnings that fall within the three respective ranges, and then adds the resulting products together. These ranges are determined by the two *bend points.* Since the cap sets the upper bound on the average lifetime earnings, it also sets the upper bound on the amount of the pension benefit.

Figure 1.1 plots the resulting schedule of replacement rates. The replacement rate is the ratio of the pension benefit to the average lifetime earnings. The two bend points are $557 and $3,362 (in terms of 2011). To determine the cap, I divide the annual maximum taxable earnings of $70,977 by 12, which equals to $5,914. As an example, consider a worker, whose average lifetime earnings are $4,000 per month. Her replacement rate is ca. 37 percent, or $1,495 per month. The pension benefit in this example is computed as follows: $0.9 \times 557 + 0.32 \times (3,362 - 557) + 0.15 \times (4,000 - 3,362)$.

---

[3]The procedure to compute the AIME is described in more detail in footnote 6 below).

The 90, 32, and 15 percent factors stayed unchanged since adoption of the 1977 Amendments. The Amendments allowed for automatic adjustment of the bend points and the cap to account for inflation; however, their value in real terms also stayed unchanged.

1.2.2 Data fact 2: Inequality in average lifetime earnings

In this section I discuss the evolution of the average lifetime earnings in the U.S. I use the 2006 Earnings Public-Use File by the SSA.[4] The data consist of a 1 percent random sample of all social security numbers issued prior to January 1, 2007. There are roughly $60, 3$ million earnings records with positive earnings values for ca. $3, 1$ million individuals who had positive earnings for at least one year during 1951 to 2006. It is a longitudinal balanced panel, since samples are selected based on the same social security number pattern every year. I merge this data set with a publicly accessible demographic subfile. This subfile provides data on the gender and the year of birth of all workers in the main data set.

There are three important caveats of these data. First, the earnings records are capped at the maximum taxable earnings threshold for a given year. Nevertheless, an average of roughly 83 percent of covered earnings have been subject to the social security tax. Second, the data don't contain information on the age of retirement of a given individual. I will have to assume that an individual retires at the normal retirement age stipulated by law for her birth cohort. Third, there is no information, whether an individual deceased during her working career. Since mortality for workers below retirement age is fairly low in the U.S., I will assume that every individual survives up to the normal retirement age. Despite these restrictions, the results in this section are qualitatively similar to Kopczuk et al. (2010), who have access to the entire database of the SSA. These data are not publicly available. In the calibration and simulation exercises, I will need to produce moments, which Kopczuk et al. (2010) do not report in their paper. This is why I will work below with the publicly available sample of the data.

I restrict attention to male workers. Women's labor force participation has been steadily increasing since the mid-1950s, which was the main driving force of inequality in average lifetime earnings among females. Including females into the sample reduces the overall inequality in average lifetime earnings (Kopczuk et al. (2010), section $V.B.$ "Cohort-Based Long-Term Inequality and Mobility").[5]

I first compute the average lifetime earnings for every worker in the data set using

---

[4]The data are available at http://www.ssa.gov/policy/docs/microdata/epuf/index.html.

[5]The authors break down the earnings of each worker by three age spans: early career (age 25 to 36), mid-career (age 37 to 48) and late career (age 49 to 60). The authors use the Gini coefficient to measure inequality in lifetime earnings. Kopczuk et al. (2010) find that the profiles of lifetime earnings inequality at each age span is steeper for male workers than for males and females together (they don't show separate profiles for females).
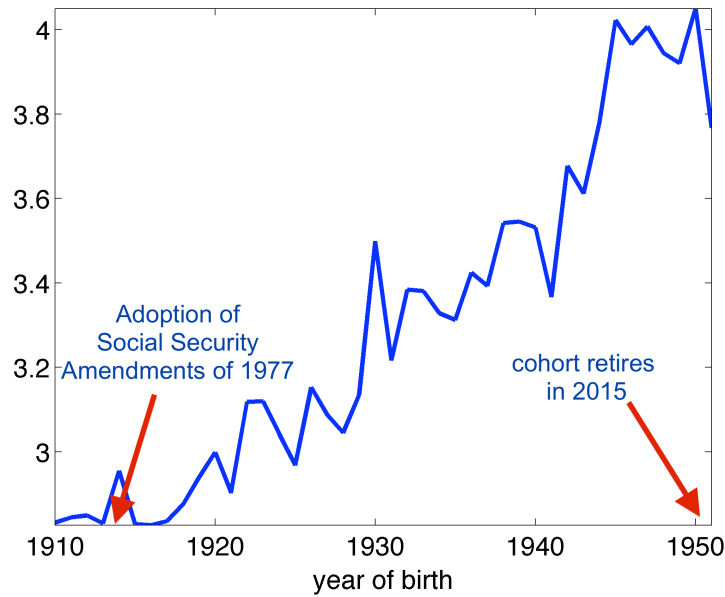
Figure 1.2: Inequality in average lifetime earnings in the data

exactly the same procedure as the one applied by the SSA.[6] Then I sort the workers by the year of birth and compute the ratio of median average lifetime earnings ("$P50$") to the average lifetime earnings of the richest household in the bottom quintile of the average lifetime earnings distribution ("$P20$").[7]

Figure 1.2 plots the results. Inequality in lifetime earnings has been rising substantially starting from the cohort born around 1915 – the $P50/P20$ percentile ratio is ca. 2.8. The inequality then gradually increases and reaches ca. 3.8 for the cohort born in 1951. This is also the last cohort, for which the data on the entire path of earnings are available.

1.3 The model

The model is based on Huggett (1996), which is a general equilibrium overlapping generations model with production, incomplete financial markets and idiosyncratic labor

---

[6]To compute average lifetime earnings per worker, I first index annual earnings of an individual using the national average wage index available at http://www.ssa.gov/oact/cola/AWI.html. Consistent with the procedure applied by the SSA, I index all earnings of each individual to the year, when she reaches the age of 62; earnings attained afterwards are not indexed. Then I compute the average of the indexed annual earnings as follows. For each worker, I take the highest n years of earnings after 1950, where n depends on the year of birth of the individual. For retirees, n equals the number of years elapsed after 1955 (or year attaining age 26, if later) and before the year attaining age 62. For instance, a retiree reaching age 63 in 1979 has $n = 22$ (the number of years between 1955 and 1978, exclusive), while a retiree reaching age 63 in 2006 has $n = 35$. If a worker has fewer than n records of earnings in the data set, I use zeros for the number of years less than n. If a worker has more than n years of earnings, then only those years with the highest annual earnings are considered.

[7]The results are qualitatively the same if one uses the $P80/P50$ ratio, the Gini coefficient, or the coefficient of variation.

productivity risk. I depart from this environment in two ways. First, I introduce earnings-dependent pension benefits. Second, I endogenize the social security policy.

1.3.1 Demographics and endowments

The economy is populated by overlapping generations of households. Each period a new generation of agents is born. The birth rate is constant and equals n. Each generation lives for J periods. Age is denoted by $j \in \{1, 2, .., J\}$. Agents enter the economy and start working at age $j = 1$. The mandatory retirement age is J^R. Agents die with probability 1 at age J.

Each agent is endowed with one unit of productive time in each period, which she supplies inelastically to a competitive labor market. Agents are born with zero assets but can accumulate savings over time, supplying capital to a competitive capital market.

At birth, each individual receives a realization of a random variable $z \in Z = \{H, L\}$, where H stands for high and L – for low ability. Abilities are drawn from a unique stationary distribution λ_z. The ability remains constant during the entire working stage of the agent. The ability has two effects. First, it affects agent's life expectancy. Denote ψ_{zj} the probability that type-z agent survives up to age j, conditional on surviving up to age $j - 1$. Second, the ability determines agent's labor productivity during the working stage.

The relative size of agents with ability z at age j can be obtained recursively as follows:
$$\mu_{z,i+1} = \psi_{zj} \frac{\mu_{zi}}{1 + n}, \text{ for } i = 1, ..., J - 1$$
with $\mu_{z,1} = \lambda_z$. At any point in time, the total measure of households is normalized to 1.

1.3.2 Labor productivity process

The productivity of type-z agent at age $j \in \{1, ..., J^R - 1\}$ is given by $\zeta_{zj} \times exp(y_{jt})$. The first term, ζ_{zj}, is a deterministic component; it captures the returns to experience over the life-cycle shared by each ability group. The second term, y_{jt}, is a random individual-specific component of log labor productivity, which follows the same stochastic process for both ability groups. It is composed of a persistent auto-regressive shock and a transitory shock:

$$
\begin{aligned}
y_{jt} &= \eta_{jt} + v_t, \\
\eta_{jt} &= \rho \eta_{j-1,t-1} + \gamma_t \text{ with } \eta_0 = 0,
\end{aligned}
\tag{1.1}
$$

where $v_t \sim \mathcal{N}(0, \sigma_{v_t}^2)$ and $\gamma_t \sim \mathcal{N}(0, \sigma_{\gamma_t}^2)$. The auto-regressive specification for η captures mean-reverting shocks, such as human capital innovations that depreciate over the life-cycle. The transitory component v represents short-term variations in individual productivity. To simplify notation below, I stack the realizations of η and v into a vector $\boldsymbol{y} \in \mathcal{Y}$ and denote agent's total efficiency units per unit of raw labor by $\epsilon_{zjt}(\boldsymbol{y})$.

1.3.3 The households

A worker supplies one unit of labor to the labor market and receive gross earnings equal to

$$e_{zjt} = w_t \epsilon_{zjt}(\boldsymbol{y})$$

with w – a competitive wage rate per efficiency unit of labor. Then, the earnings net of the social security taxes amount to:

$$e_{zjt} - \tau_t \min(cap_t, e_{zjt}),$$

where the linear social security tax rate, τ, applies to the portion of gross earnings below the maximum taxable earnings, cap.

Agent's average lifetime earnings evolve according to:

$$\bar{e}_{z,j+1,t+1} = \begin{cases} \frac{1}{j}\left[(j-1)\bar{e}_{zjt} + \min(cap_t, e_{zjt})\right] & \text{for } 2 \le j < J^R; \\ \bar{e}_{zjt} & \text{for } j \ge J^R. \end{cases} \tag{1.2}$$

with $\bar{e}_{z1t} = 0$.

When the agent with average lifetime earnings $\bar{e}_{zjt}$ retires, she receives a pension benefit, $B_{SS}(\bar{e}_{zjt})$, given by

$$B_{SS}(\bar{e}_{zjt}) = \begin{cases} \alpha_1 b^{\alpha_2 - 1} \bar{e}_{zjt} & \text{for } \bar{e}_{zjt} < b \\ \alpha_1 \bar{e}_{zjt}^{\alpha_2} & \text{for } \bar{e}_{zjt} > b, \end{cases} \tag{1.3}$$

with $\alpha_1 \in \mathcal{R}_+$, $\alpha_2 \in \mathcal{R}$ and $b \in \mathcal{R}_+$. The first line of the pension benefit schedule formalizes the idea that the replacement rate for agents with lifetime earnings below the bend point b is constant and equal to $\alpha_{1,t} b^{\alpha_{2,t}-1}$. The second line shows the pension benefit for agents with lifetime earnings above b. I will show in the calibration section that the chosen functional form fits fairly accurately the replacement rates implied by the pension benefit formula in the U.S.

Figure 1.3 plots the replacement rate as a function of average lifetime earnings for different values of α_2 (with $b = 0.5$). For $\alpha_2 < 1$, the replacement rate (weakly) decreases with average lifetime earnings. As a special case, the pension benefits are lump-sum if
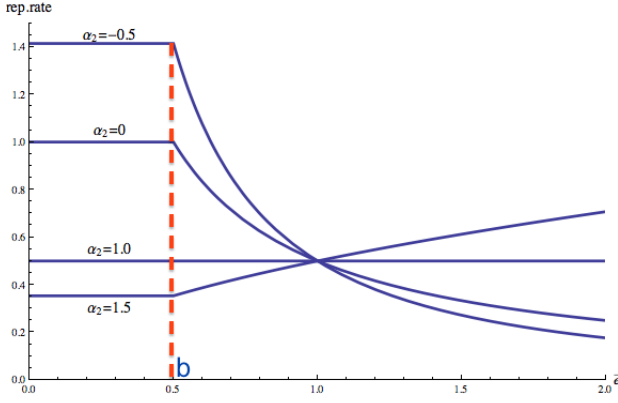
Figure 1.3: Replacement rate as a function of α_2
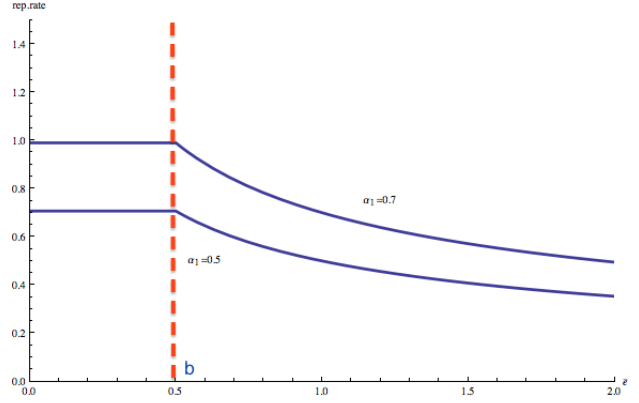
Figure 1.4: Replacement rate as a function of α_1

$\alpha_2 = 0$ and $b = 0$. If $\alpha_2 = 1$, every individual receives the same fraction of her lifetime earnings as a pension benefit. If $\alpha_2 > 1$, the replacement rate turns into a (weakly) increasing function of lifetime earnings. One can see from this figure the importance of the bend point b: it prohibits the replacement rates for low earners from being infinite. Figure 1.4 plots the replacement rate as a function of average lifetime earnings for different values of α_1 (again, $b = 0.5$). The variable α_1 is the benefit adjustment factor: it scales the benefits for all individuals by shifting up and down the replacement rate curve. Another way to think about α_1 is to observe that it equals the replacement rate for the agent with lifetime earnings equal to 1.0 (as long as $b < 1.0$).

I can now write down agent's problem using the dynamic programming notation. Let $\boldsymbol{x}$ denote the individual state of the agent in period t:

$$\boldsymbol{x} = (z, j, \boldsymbol{y}, a, \bar{e}),$$

where $a \in A = [0, a^{max}]$ – asset holdings and $\bar{e} \in \bar{E} = [0, \bar{e}^{max}]$ – average lifetime earnings.[8] Furthermore, denote $F_t(\boldsymbol{x})$ the cumulative population density function of agents at the beginning of period t. Let $\boldsymbol{\Psi}_t$ be the government social security policy at the beginning of period t:

$$\boldsymbol{\Psi}_t \equiv \{\tau_s, \alpha_1, \alpha_2\}_{s=t}^\infty,$$

where I assume that the bend point b is a parameter. Finally, let $V(\boldsymbol{x}; \boldsymbol{\Psi}_t, F_t)$ be the discounted lifetime indirect utility of the agent in state $\boldsymbol{x}$ at time t.

Taking the government policy $\boldsymbol{\Psi}_t$ as given, agents solve:

$$V(\boldsymbol{x}; \boldsymbol{\Psi}_t, F_t) = \max_{c, a'} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta \psi_{z,j} \mathbb{E}\left[V(\boldsymbol{x}'; \boldsymbol{\Psi}_{t+1}, F_{t+1}) \mid \boldsymbol{x}\right] \right\} \tag{1.4}$$

---

[8]Since there will be only one type of asset in the economy, will I refer to a as capital, wealth and assets interchangeably.

subject to the laws of motions:

$$\boldsymbol{x}' = (z, j+1, \boldsymbol{y}', a', \bar{e}') \tag{1.5}$$

$$a' = \begin{cases} (1+r_t)a + e_{zjt} - \tau_t \min(cap_t, e_{zjt}) + T_{LS,t} + T_{W,t} \\ -\tau_{M,t} \min(cap_{M,t}, e_{zjt}) - \tau_I(r_t a + e_{zjt}) - c & \text{if } j < J^R \\ (1+r_t)a + B_{SS,t}(\bar{e}_{zjt}) - \tau_{I,t}(r_t a + B_{SS,t}(\bar{e}_{zjt})) + T_{LS,t} + T_{R,t} + T_{M,t} - c & \text{if } j \geq J^R. \end{cases} \tag{1.6}$$

$$F_{t+1} = G(F_t) \tag{1.7}$$

including the law of motion for the average lifetime earnings in eq. (1.2), and the non-negativity constraints:

$$c \geq 0 \text{ and } a' \geq 0. \tag{1.8}$$

In eq. (1.4), c denotes consumption and γ controls the degree of relative risk aversion; β is the subjective discount factor and $\mathbb{E}$ is a conditional expectation operator. Eq. (1.5) is a law of motion for the individual state; note that the ability z doesn't change during agent's life. Eq. (1.6) is a budget constraint of a working agent ($j < J^R$) and a retired agent ($j \geq J^R$). A working agent receives earnings net of the social security contributions and a gross interest on savings from the previous period, $(1+r)a$, where r denotes the interest rate. The worker pays a linear hospital insurance tax, τ_M, on the portion of earnings below the maximum taxable earnings for hospital insurance, cap_M. Besides, the agent receives a lump-sum transfer, T_{LS}, a means-tested transfer for working households, T_W, which equals to zero if agent is not eligible, and pays income taxes on gross earnings and earned interest according to the function, τ_I. A retired agent receives a pension benefit, B_{SS}, a gross interest on savings, a lump-sum income transfer, a means-tested transfer for retired households, T_R, and a lump-sum Medicare transfer, T_M, which equals to zero if agent is not eligible. The earned interest and the pension benefit are subject to income taxes.

The perceived law of motion for the distribution of agents over the states is denoted by G in eq. (1.7). I exclude agents from borrowing, which explains the non-negativity constraint on assets in eq. (1.8).

The solution to the household's problem generates decision rules $c(\boldsymbol{x}; \boldsymbol{\Psi}_t, F_t)$ and $k'(\boldsymbol{x}; \boldsymbol{\Psi}_t, F_t)$.

1.3.4 The government

The government is involved in three activities. First, it runs a pay-as-you-go social security system: it collects payroll contributions from workers and redistributes them among retirees. Second, the government provides means-tested transfers to workers, T_W, and retirees, T_R, as well as lump-sum benefits to all individuals, T_{LS}. To finance the benefits,

the government taxes earnings, capital interest and pensions based on the tax function τ_I and confiscates the wealth left by deceased agents. Third, the government runs the Medicare program: it collects hospital insurance taxes from workers and redistributes them equally among retirees. The government is assumed to run a balanced budget in each of these activities.[9]

1.3.5   The firm

The aggregate output technology is represented by a standard Cobb-Douglas production function $K^{\theta} N^{1-\theta}$, where K is the aggregate capital stock, N – aggregate effective labor, and $\theta \in (0,1)$ – capital share in production. The output can be consumed or invested in capital. The depreciation rate of capital is $\delta \in (0,1)$. The firm produces output goods and sells them in a competitive market at a price that is normalized to one. As standard with a constant returns to scale technology and perfect competition, I assume that there exists a representative firm, which operates this technology. This implies that the factor prices are given by:

$$r_t = \theta K_t^{\theta-1} N_t^{1-\theta} - \delta \text{ and } w_t = (1-\theta) K_t^{\theta} N_t^{-\theta}. \tag{1.9}$$

1.3.6   Recursive competitive equilibrium

I define the recursive competitive equilibrium in two steps. In the first step, I set up a recursive competitive equilibrium, when agents take the policy vector, $\boldsymbol{\Psi}_t$, as given. In the second step, I make $\boldsymbol{\Psi}_t$ consistent with optimization problem of the social planner.

Definition. A recursive competitive equilibrium with an exogenous social security policy consists of the sequence of value functions, $\{V(\boldsymbol{x}; \boldsymbol{\Psi}_s, F_s)\}_{s=t}^{\infty}$ and individual decision rules $\{c(\boldsymbol{x}; \boldsymbol{\Psi}_s, F_s), \ a'(\boldsymbol{x}; \boldsymbol{\Psi}_s, F_s)\}_{s=t}^{\infty}$, the sequence of prices $\{w_s, r_s\}_{s=t}^{\infty}$ and transfers $\{T_{W,s}, T_{R,s}, T_{LS,s}\}_{s=t}^{\infty}$, and the sequence of distributions $\{F_s(\boldsymbol{x})\}_{s=t}^{\infty}$, such that *given* $\boldsymbol{\Psi}_s$ the following statements hold for all $s = t, \ldots, \infty$:

- given $(\boldsymbol{\Psi}_s, T_{LS,s}, T_{W,s}, T_{R,s}, r_s, w_s, G)$, the functions $c(\boldsymbol{x}; \boldsymbol{\Psi}_s, F_s)$ and $a'(\boldsymbol{x}; \boldsymbol{\Psi}_s, F_s)$ solve agent's optimization problem in eq. (1.4);

- the social security budget is balanced according to :

$$\sum_{j=1}^{J^R-1} \int_{A \times \bar{E} \times \mathcal{Y}} \tau_t \min(cap_t, e_{zjt}) dF_t(\boldsymbol{x}) = \sum_{j=J^R}^{J} \int_{A \times \bar{E} \times \mathcal{Y}} B_{SS,t}(\bar{e}_{zjt}) dF_t(\boldsymbol{x}); \tag{1.10}$$

---

[9]At the time period, for which I calibrate the benchmark model (1977), the Social Security Trust Fund was indeed running a balanced budget with expenditures roughly equal to contributions, see http://www.ssa.gov/oact/STATS/table4a3.html. During 1978-2008, the system was generating surpluses, which dried up in 2009. Since 2009, the system has been running steadily growing deficits.

- the Medicare and the income transfer program each run a balanced budget:

$$\sum_{j=1}^{J^R-1} \int_{A\times\bar{E}\times\mathcal{Y}} \tau_M \min(cap_{M,s}, e_{zjs})dF_t(\boldsymbol{x}) = T_{M,s} \sum_{j=J^R}^{J} \sum_{z} \mu_{z,j}; \qquad (1.11)$$

$$\sum_{j=1}^{J} \int_{A\times\bar{E}\times\mathcal{Y}} \tau_{I,s}(\iota_s)dF_s(\boldsymbol{x}) \;+\; \sum_{j=1}^{J} \int_{A\times\bar{E}\times\mathcal{Y}} (1-\psi_{z,j})a'(\boldsymbol{x};\boldsymbol{\Psi}_s, F_t)dF_s(\boldsymbol{x}) \qquad (1.12)$$

$$= \sum_{j=1}^{J^R-1} \int_{A\times\bar{E}\times\mathcal{Y}} T_{W,s}dF_s(\boldsymbol{x}) + \sum_{j=J^R}^{J} \int_{A\times\bar{E}\times\mathcal{Y}} T_{R,s}dF_s(\boldsymbol{x})$$

$$+ \sum_{j=1}^{J} \int_{A\times\bar{E}\times\mathcal{Y}} T_{LS,s}dF_s(\boldsymbol{x})$$

- given K_s and N_s, the factor prices r_s and w_s are determined by eq. (1.9);

- the capital and labor markets clear, i.e.:

$$K_s \;=\; \sum_{j=1}^{J} \int_{A\times\bar{E}\times\mathcal{Y}} k(\boldsymbol{x};\boldsymbol{\Psi}_s, F_s)dF_s(\boldsymbol{x}) \text{ and } N_s = \sum_{j=1}^{J^R-1} \sum_{z} \mu_{zjs}\epsilon_{zjs}(\boldsymbol{y});$$

Definition. A **steady-state recursive competitive equilibrium with an exogenous social security policy** is a recursive competitive equilibrium with $F_{t+1} = F_t$ for all t.

1.3.7 Social planner's problem

In this section I show how the social security policy, $\boldsymbol{\Psi}$, arises endogenously. I use a social welfare approach to endogenize the social security policy: a utilitarian government maximizes the weighted sum of the expected discounted lifetime utilities of all generations, who are alive in the period, when the change to social security policy is made. I assume that the Pareto weight is a function of agent's average lifetime earnings and age. I discuss this assumption in the next section. I furthermore assume that the maximum taxable earnings threshold, *cap*, and the bend point, *b*, are exogenous. I make this assumption to overcome the computational burden but this doesn't mean that the earnings cap and the bend point are unimportant policy variables. I discuss how the results are likely to change, once one allows the government to choose these two variables, as well.

The government sets the social security policy once-and-for-all. Formally, the government solves in period t:

$$\boldsymbol{\Psi}_t^{\star} = \arg\max_{\boldsymbol{\Psi}_t} \sum_{j=1}^{J} \int_{A \times \bar{E} \times \mathcal{Y}} \omega(\bar{e}_{zjt}, j) V(\boldsymbol{x}; \boldsymbol{\Psi}_t, F_t) dF_t, \tag{1.13}$$

subject to eq. (1.10), (1.11), and (1.12), where $\omega(\cdot)$ is a Pareto weight function (to be specified below) of agent's lifetime earnings, $\bar{e}$, and age, j. More precisely, in period t the government sets (α_1, α_2), while the tax rate, τ, adjusts in each period to satisfy the government budget constraint in eq. (1.10).The government chooses the policy under rational expectations about the effects of these policies on future equilibrium outcome and the welfare of each agent during her entire lifetime.

The government faces non-trivial tradeoffs when choosing the parameters of the pension benefit formula. An increase in α_1 directly raises the total size of the pension fund and for a given α_2 increases the amount of redistribution. By choosing α_2, the government can direct income redistribution towards the agents with lower lifetime earnings because they have higher marginal utility of consumption. Furthermore, in the absence of formal or informal private insurance markets against idiosyncratic uncertainty, the public pension system provides a partial substitute for these missing markets and therefore may lead to less volatile household consumption over time. Combining the two policies α_1 and α_2 together, the government can achieve a more equal distribution of income, and therefore, wealth, consumption and welfare. But redistribution and insurance go hand-in-hand with distortions. A larger size of the system has to be financed through higher social security taxes, which reduce agent's individual and aggregate saving. Similarly, for a given size of the pension fund, targeting the pension benefit at specific groups distorts the savings of these groups because they anticipate a higher future transfer in retirement. Due to the distortionary effects, the pension system also affects prices of labor and capital and the government lump-sum transfers through the general equilibrium.

Definition. A **steady state recursive competitive equilibrium with an endogenous social security policy** is

- a set of functions $(T_{LS,s}, T_{W,s}, T_{R,s}, r_s, w_s, G, c, a')$, which, given $\boldsymbol{\Psi}_s$, satisfy the definition of the recursive competitive equilibrium;

- a function $\omega(\bar{e}_{zjt}, j)$, such that $\boldsymbol{\Psi}_s = \boldsymbol{\Psi}_t^{\star}$ for all $s = t, \ldots, \infty$.

1.4 Calibration of the benchmark model

The benchmark model economy is calibrated to the U.S. data in 1977. First, I calibrate the parameters of the benchmark model, which describe the economic environment, so that the model endogenously matches the key moments of the lifetime earnings distribution during that period. Then I calibrate the political power of the old-age population in such

Parameter	Description	Value	Source
- Demographics:			
J	max. life span	65	
J^R	retirement age	46	
n	birth rates	1.80 %	World Bank, 1979
$\psi_{z,j}$	cond. survival prob.	vectors	Elo & Preston (1996)
- Preferences and technology:			
γ	degree of risk aversion	2.0	fixed
θ	capital share	0.36	Cooley & Prescott (1995)
- Income tax and transfer systems:			
m_0	average tax rate	0.48	1979, Gouveia & Strauss (1994)
m_1	progressivity	0.82	1979, Gouveia & Strauss (1994)
$\underline{c}_W$	cons. floor workers	$14,515	1980, UKCPR
$\underline{c}_R$	cons. floor retired	$10,067	1980, UKCPR
- Social security:			
cap	earnings cap	$61,256	SSA, 1977
b	bend point	$6,695	SSA, 1977
α_1	level	0.514	estimated by NLLS, 1977
α_2	curvature	0.577	estimated by NLLS, 1977
- Medicare:			
τ_M	hospital insurance	1.8 %	SSA, 1977
cap_M	earnings cap	$61,256	SSA, 1977
- Labor productivity:			
λ_H	share college degree	22 %	CPS 1977
$\zeta_{j,z}$	age-efficiency profile	vectors	CPS, 1977
ρ	AR(1) coefficient	0.97	Heathcote et al. (2010)
σ_v^2	var. transitory shock	0.056	Heathcote et al. (2010), 1977

Table 1.1: Parameters of the benchmark model obtained outside the model

a way, that the model endogenously produces the pension benefit schedule, which the Congress adopted at that time.

The parameters of the model can be grouped into three different sets: (i) demographics $\{J, J^R, n, \psi_{z,j}\}$; (ii) preferences and technology $\{\gamma, \beta, \theta, \delta\}$; (iii) productivity parameters $\{\zeta_{j,z}, \sigma_v^2, \sigma_\gamma^2, \rho, \lambda_z\}$; (iv) government parameters, which build four subgroups: the parameters of the social security policy, $\{\tau, \alpha_1, \alpha_2, cap, b\}$, the Medicare program, $\{\tau_M, cap_M\}$, the income tax and transfer systems, $\{\tau_I, \underline{c}_W, \underline{c}_R\}$, and the Pareto weight function, $\omega(\cdot)$.

I proceed as follows. First, I parametrise the model to match the key target moments from the U.S. data on the evolution of lifetime earnings inequality, while holding the social security variables exogenously given. Second, I specify and calibrate the Pareto weight function, so that the social security policies arise endogenously in a political equilibrium. Table 1.1 shows all the parameter values obtained outside the model; table 1.2 shows the parameters calibrated using the model. All dollar amounts in this section are in 2011 dollars.

Parameter		Value	Target
- Preferences and technology:			
β	discount factor	1.01	$K/GDP = 3.30$
δ	depreciation	6 %	$investment/GDP = 6$ %
- Income taxes:			
m_2	scale	0.13	$transfers/GDP = 14$ %
- Labor productivity:			
σ_γ^2	var. persistent shock	0.04	P50/P20 lifetime earnings
- Social planner's problem:			
κ_1	earnings bias	-2.16	$\alpha_{1,1977}^{data}$
κ_2	age bias	4.27	$\alpha_{2,1977}^{data}$

Table 1.2: Calibrated parameters of the model

1.4.1 Calibration for given policies

One model period equals to one year. A high ability type in the model corresponds to a head of the household with at least 16 years of schooling in the data; a low ability type is a household with fewer years of schooling.

An agent in the model corresponds to a household in the data. The key data source for the calibration is CPS. I restrict attention to the households with a male head of working age (20-64), who works at least 260 hours per year. Household's earnings and income are adjusted using the OECD equivalence scale.

1.4.1.1 Demographics

The maximum possible age, J, is set to 65 periods, which corresponds to real age 85. The retirement age, J^R, is 45, which corresponds to real age 65. This is the retirement age, at which retirees born in 1939 and earlier are entitled to a full pension benefit. The birth rate, n, equals 1.80 percent, which is the average birth rate over the period 1960-77.[10] I take the series of conditional survival probabilities for males for 1979-81 from Elo and Preston (1996).

1.4.1.2 Preferences and technology

The discount factor, β, is calibrated to match the capital to output ratio of around 3.3, as reported in Cooley and Prescott (1995). The degree of risk aversion, γ, is set to 4.0, so that the elasticity of inter-temporal substitution of consumption is 0.25. The capital share, θ, is chosen to match the labor share of 64 percent, while the depreciation rate of capital, δ, is calibrated to match the investment to capital ratio of ca. 0.06; these values are consistent with Cooley and Prescott (1995).

---

[10]World Development Indicators, World Bank, http://data.worldbank.org/indicator/SP.POP.GROW?page=5.

1.4.1.3   Government parameters

Social security

The political element of the model is not exploited at this stage of the calibration, since I treat social security as exogenous.

Using the non-linear least squares estimator, I estimate the parameters α_1 and α_2 of the pension benefit function using the statutory replacement rates from figure 1.1 for the portion of lifetime earnings above the first bend point (which in 1977 equaled \$6,694). Since I change the units of measurement for income in the model, the estimate of α_1 has to be re-adjusted to $s^{\hat{\alpha}_2-1} \times \hat{\alpha}_1$, where s is the ratio of data units to model units.[11]

The annual maximum taxable earnings threshold, *cap*, was \$61,256 in 1977.[12]

Income tax and transfer systems

I model the means-tested transfers to workers, T_W, and retirees, T_R, using the specification in R. Glenn Hubbard (1995). Working households receive a transfer given by:

$$T_{W,t} = \max\left\{0, \underline{c}_{W,t} - [e_{zjt} + (1+r_t)a_t]\right\},$$

while retired households receive a transfer:

$$T_{R,t} = \max\left\{0, \underline{c}_{R,t} - [B_{SS,t}(\bar{e}_{zjt}) + (1+r_t)a_t]\right\}.$$

These transfer functions guarantee a pre-tax income of $\underline{c}_W$ and $\underline{c}_R$ to workers and retirees, respectively. To set the two consumption floors, I use the UKCPR National Welfare Data.[13] These data are at the state level, so I weight expenditures by each state's population.

The individual income tax function follows Gouveia and Strauss (1994). It is a commonly used specification in the empirical macroeconomic literature. This progressive taxation rule reads:

$$\tau_{I,t}(\iota_t) = m_{0,t}[\iota_t - (\iota_t^{-m_1} + m_2)^{-1/m_1}],$$

where $\tau_{I,t}(\iota)$ is the amount of taxes the agent has to pay if her taxable income equals ι at time t. For workers, the taxable income is given by the sum of interest rate payments and earnings before social security taxes, i.e. $\iota_t = r_t a_t + e_{zjt}$; for retirees, the taxable

---

[11]To establish the relationship between the model units and the data units, I use the average equalized household earnings in CPS in 1977 (in terms of 2011 dollars).

[12]The dollar value for the cap has to be converted to model units.

[13]http://www.ukcpr.org/data. The benefits for working households include Aid to Families with Dependent Children (AFDC), Temporary Assistance for Needy Families (TANF) and Food Stamps (SNAP). The benefits for retired households include Supplemental Security Income and Food Stamps.

income is given by the sum of interest rate payments and social security benefits, i.e. $\iota_t = r_t a_t + B_{SS,t}(\bar{e}_{zjt})$.[14]

In the schedule above, the parameter $m_0 \in [0,1]$ is the marginal (and average) tax rate as income goes to infinity. The parameter $m_1 \in [-1,\infty)$ determines the curvature of the marginal tax function $\tau_I'(\iota)$. m_2 is a scaling parameter. I set m_0 to 0.479 and m_1 to 0.817, which are the estimates obtained by Gouveia and Strauss (1994) for 1977. I calibrate m_2 to match the share of government transfers in GDP in 1979.[15]

Medicare

The employer's and the employee's contribution to the hospital insurance was 0.9 percent in 1977, so I set τ_M to 1.8 percent. Consistent with the data, the maximum taxable earnings threshold for the hospital insurance, cap_M, was the same as for social security.

1.4.1.4   Labor productivity

A household's labor productivity depends on three components: a deterministic ability dependent age-efficiency profile, ζ, a transitory shock, v, and a persistent shock, η.

The values for the variance of the transitory shock, σ_v^2, and the autoregressive coefficient, ρ, are taken from Heathcote et al. (2010b) with $\rho = 0.97$ and $\sigma_{v,1977}^2 = 0.056$.[16] Then I construct an i.i.d. two-state Markov chain with equal probabilities. I calibrate the variance of the persistent shock, σ_γ^2, to match the $P50/P20$ ratio of the distribution of lifetime earnings in the data. This moment is highly informative of the parameter σ_γ^2 because lower values of the parameter are associated with a lower dispersion in cross-sectional earnings and – given that the process is highly persistent – lower inequality in average earnings over the lifetime, while the opposite is true for higher values of σ_γ^2.

---

[14]According to the U.S. law, payroll taxes withheld from earnings are not deductible from federal or any state income tax. Furthermore, pension benefits are subject to income taxation since 1984.

[15]I use the data from "Historical Effective Tax Rates, 1979 to 2005" (Table 5. "Total Income and Total Federal Tax Liabilities for All Households, by Household Income Category, 1979 to 2005") by the Congressional Budget Office, which are available at http://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/98xx/doc9884/12-23-effectivetaxrates_letter.pdf. To find the aggregate amount of transfers, I sum up cash transfers (excluding pensions) and in-kind income. To find the aggregate output, I sum up pre-tax wages, proprietor's income, other business income, interest and dividends, and other income. I exclude capital gains because they are not part of the model.

[16]The authors produce annual estimates of these parameters using PSID for 1967-2000. The authors restrict attention to married households, in which the husband is between 25 and 59 years old and works at least 260 hours per year. These sample selection criteria are different from the ones I apply in CPS. In particular, my sample includes not only married but also single households, as long as the head is a male. Also, I include the households, whose head is between 20 and 64 years old. This discrepancy in sample selection criteria can potentially be problematic. However, Guvenen (2009), who uses PSID for 1968-1993 applying the same selection criteria as me, obtains annual estimates, which are very similar to the ones by Heathcote et al. (2010b). I prefer to use the estimates by Heathcote et al. (2010b) because of a larger time span.

In figure 1.2, I showed the evolution of $P50/P20$ ratio of the distribution of lifetime earnings for workers across 1910-51 cohorts. This ratio was constructed using the same procedure as the one applied by the SSA (see footnote 6 for details). Since I cannot treat the earnings in the same way in the model, the model distribution and the empirical distribution behind figure 1.2 would not be comparable. Therefore, I re-compute the average lifetime earnings for each worker in the data as follows. First, I compute lifetime earnings inequality for workers at age 46. By law, the SSA computes the average of earnings earned after 1951, and this is the first available year in the data set by the SSA. In 1977, the workers of age 46 build the first cohort, for which I have a sequence of earnings records starting from 1951, when this cohort turned 20. Second, only positive earnings earned flow into the computation of the average lifetime earnings. I do so because in the CPS data, which I use at different points during calibration, I follow the sample selection strategy by Heathcote et al. (2010a) and drop those individuals who work fewer than 260 hours per year. Second, the average is computed over *all* positive earnings and not over a certain number of years with *highest* earnings. Finally, I index all earnings of all workers to the year 1977, rather than indexing the earnings of each worker to the year, in which she reaches age 62. After constructing the average lifetime earnings for each worker in the way described above, I compute the $P50/P20$ ratio.

After computing the target $P50/P20$ ratio in the data, I calibrate the variance of the persistent shock, σ_γ^2, to match this ratio. Then I construct an age-dependent Markov chain for the autoregressive process using 4 equally-spaced nodes at each age. Given $\mu_0 = 0$, the conditional variance of η_j increases with the age.

I construct the age-efficiency profiles from the CPS.[17] The share of high ability agents, λ_H, is 0.22, which corresponds to the average share of male college graduate heads in CPS in 1977.

1.4.1.5   Model fit

Table 1.3 evaluates the performance of the calibrated benchmark model economy.[18] The model overestimates the inequality in pre-government earnings. This is not surprising, since I chose to calibrate the variance of the persistent shock, σ_γ^2, to match the inequality in lifetime earnings rather than cross-sectional earnings. This also explains slightly over-estimated inequality in pre-government incomes. However, the model fits fairly well the

---

[17]Following Hansen (1993), I compute mean hourly earnings of high school and male college graduates of age 20-64 in CPS. Then I normalize the mean hourly earnings of both types by the average mean hourly earnings across both types and all age groups in 1977. Finally, I fit a quadratic polynomial curve to obtain a smoother approximation of the two age-efficiency profiles.

[18]The moment for the earnings and income distribution are obtained using CPS for 1965-77. The Gini index for net worth is based on the household-level data from the SCF for 1987, see Quadrini et al. (1997). The social security moments are taken from the *2014 Annual Report Of The Board Of Trustees Of The Federal Old-Age And Survivors Insurance And Federal Disability Insurance Trust Funds* and computed for the year 1977.

inequality at the bottom of the pre-government income distribution. This will play an important role when it comes to calibrating the Pareto weights in the next section. As typical of these types of models, the inequality at the top of the wealth distribution is underestimated, which leads to a much lower Gini coefficient than in the data.

Moment	Model	Data
Pre-govt. earnings distribution:		
- P50/P20	2.04	1.83
Pre-govt. income distribution:		
- P50/P20	1.86	1.75
- Q1	5.75	5.47
- Q2	11.10	11.90
- Q3	19.52	16.72
- Q4	39.77	22.90
- Q5	23.85	42.96
Net worth distribution:		
- Gini index	0.56	0.78
Social security system:		
- tax rate τ, %	7.55	8.66

Table 1.3: Untargeted moments in the benchmark model

1.4.2　Calibration with endogenous policies

It is not surprising that with equal Pareto weights, the government in the model prefers a higher level of taxes and larger replacement rates for agents with low lifetime earnings than the ones observed in the real world. To be precise, the government chooses $\alpha_1 = 1.79$ and $\alpha_2 = -0.80$. Intuitively, the government in the model uses the social security system to shift consumption towards wealth-poor retirees, since they have high marginal utilities of consumption.[19] In this section, I specify and calibrate the Pareto weight function, $\omega(\cdot)$, so that the social security policy implied by the Social Security Amendments of 1977 arises endogenously as a solution to the government maximization problem in eq. (1.13). I discuss the most preferred policies of agents in section 1.4.2.1. This discussion provides intuition for how to specify the Pareto weight function and estimate its parameters (section 1.4.2.2).

1.4.2.1　Most preferred policies

Suppose first that the policy variable α_1 is fixed. In order to understand how agents' most preferred policy α_2 differs for different levels of lifetime earnings, I conduct the following experiment. I split the entire population in the model into 4 age groups: 1) early career:

---

[19]It might be for the same reason why prominent politico-economic models by Corbae et al. (2009) and Song et al. (2012) predict for the U.S. economy a way too high equilibrium level of income redistribution.

(real-life) age $20 - 36$, 2) mid-career: age $37 - 48$, 3) late career: age $49 - 64$, and 4) retirees: age $65 - 85$. Then I split each of these age groups into 7 subgroups based on the lifetime earnings of the agents in that group (from lowest to highest, where the exact range depends on the age group). Now I solve the government optimization problem in eq. (1.13) assuming that α_1 is fixed and the government puts the entire weight on one of the 28 groups.

The equilibrium outcomes in each of the experiments are presented in figure 1.5 (ignore dashed horizontal lines for now). As can be seen from the figure, the most preferred α_2 increases in average lifetime earnings for workers within age subgroups. In order to understand why, it is helpful to look at the dynamics of the aggregate variables, as the economy transits from the initial steady state to the new steady state associated with a given policy. Figure 1.7 plots the transitional dynamics of the interest rate, the social security tax rate, the wages and the transfers for selected values of α_2 (keeping α_1 constant).

Workers with low lifetime earnings prefer lower values of α_2. These workers have had bad draws of productivity or/and have low deterministic profiles of earnings due to low ability (in the benchmark model, the correlation between earnings and lifetime earnings is 0.73). Since productivity shocks are persistent, these agents expect lower profiles of earnings in the future. This is why these agents opt for a system with a high degree of intra-generational redistribution. They also value the insurance nature of the intra-generational redistribution. Furthermore, poor workers derive most of their income from earnings, so they also benefit from higher after-tax earnings due to rising pre-tax wages and falling social security tax rate. Obviously, retired agents with low lifetime earnings strongly value the redistributive nature of the social security system and therefore opt for a low α_2. These agents accept declining interest rates, since their saving are low in the first place (the correlation between lifetime earnings and asset holdings is 0.89). As lifetime earnings of workers rise, agents start to prefer a social security system with more earnings-oriented pension benefits. Also, these agents benefit from the rising interest rate. However, they have to trade off these benefits against increasing social security taxes. Since older agents have shorter spans of working career, the taxation cost is small for them, so they choose higher values of α_2.

How do agents' preferences differ across α_1? I conduct a similar experiment as before but now assume that α_2 is fixed. The equilibrium outcomes as a solution to the government maximization problem are shown in figure 1.6. Figure 1.8 plots the transitional dynamics of the interest rate, the social security tax rate, the wages and the transfers for selected values of α_1 (keeping α_2 constant). For young agents, the prospect of lower taxes on earnings for the rest of their working career outweighs the cost of having contributed for some time to the pension system without getting any benefits, so these agents prefer lower α_1. For the older agents, the reverse is true: they choose higher α_1 to sacrifice higher taxes
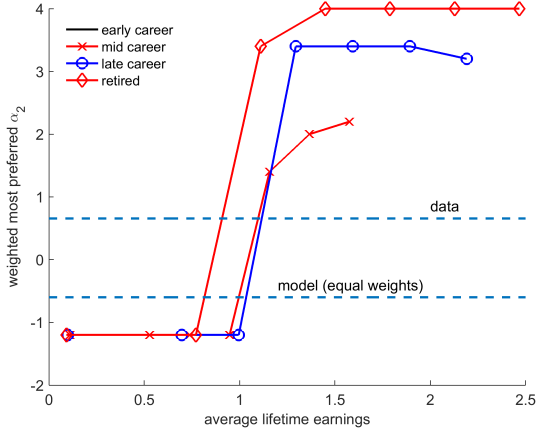
Figure 1.5: Optimal policy α_2 for different subgroups of agents (keeping α_1 at the level of the benchmark model economy)
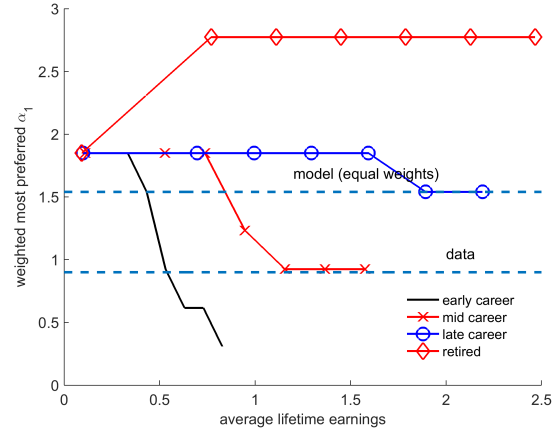
Figure 1.6: Optimal policy α_1 for different subgroups of agents (keeping α_2 at the level of the benchmark model economy)

on earnings for the benefits stemming from larger pensions and higher returns on savings. Within each working age group, α_1 decreases with lifetime earnings, since earnings-rich agents want to benefit from rising after-tax earnings. For retirees, α_1 increases in lifetime earnings due to growing pensions and interest rates.

1.4.2.2 Pareto weight function

The discussion of the most preferred policies suggests one strategy for how to specify the Pareto weight function. Return to figure 1.5. Here, the upper dashed horizontal line refers to the policy α_2 estimated for the benchmark model in 1977. The lower dashed horizontal line refers to the policy obtained from solving the government maximization problem in eq. (1.13) assuming equal Pareto weights. The latter policy is obtained by weighting the most preferred policies of all groups of agents with the weights given by the relative population sizes of each group. As can be seen, the government in the model chooses a way too low α_2 which corresponds to way too high replacement rates for agents with low lifetime earnings. This is not surprising, since these agents are typically earnings-poor and wealth-poor workers and wealth-poor retirees with very high marginal utility from consumption. In order to match the policy α_2 in the data, one could specify the Pareto weights as an increasing function of agent's lifetime earnings.

As for the policy α_1, the model overpredicts the size of the pension system (figure 1.6). One way to bring the model closer to the data in terms of the policy α_1 is to

Figure 1.7: Transitional dynamics of aggregate variables for selected values of α_2 (keeping α_1 constant)

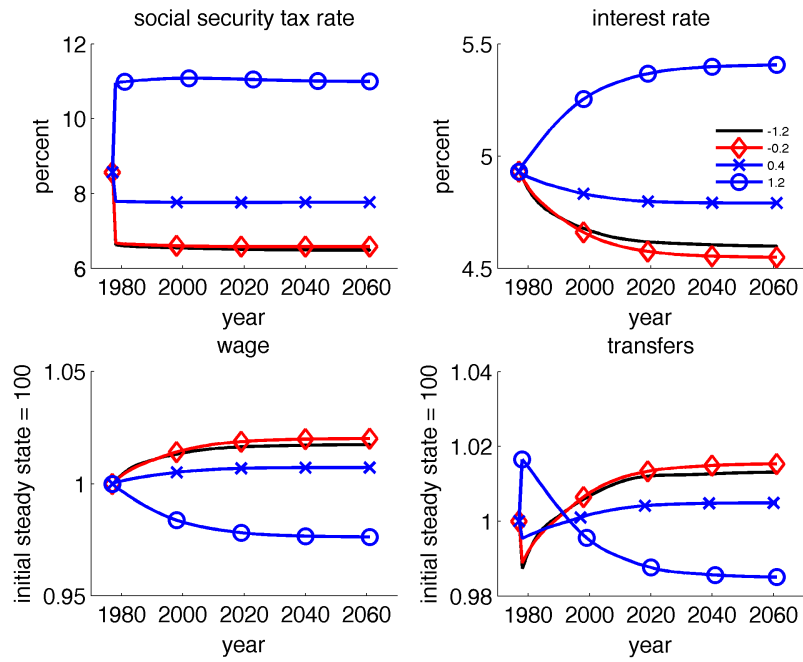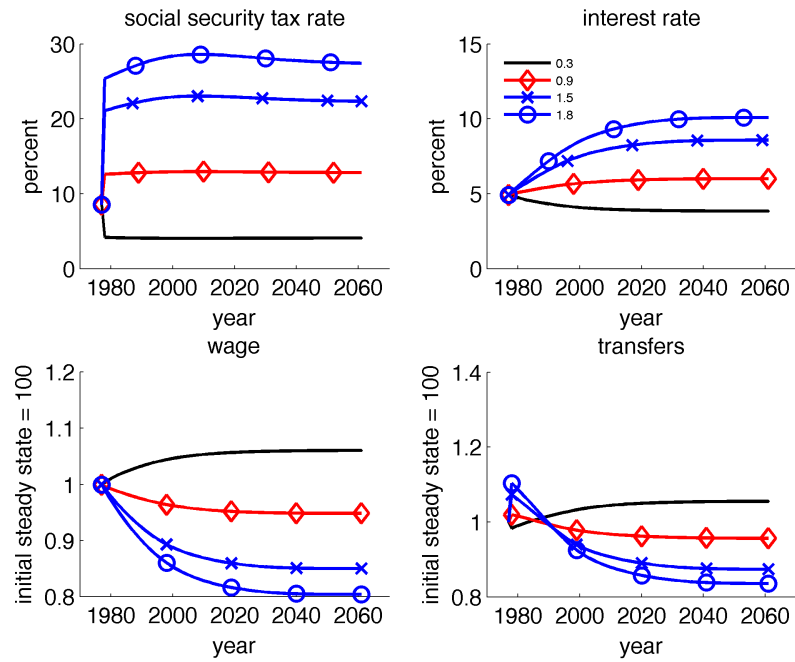Figure 1.8: Transitional dynamics of aggregate variables for selected values of α_1 (keeping α_2 constant)
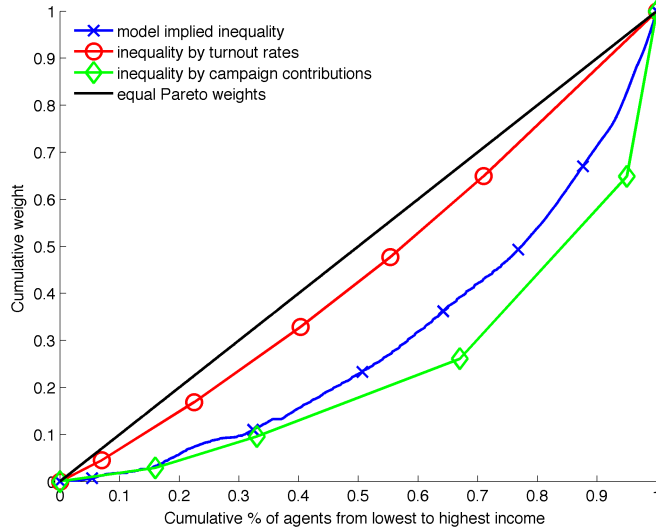
Figure 1.9: Political Lorenz curve implied by the model and the data

specify the Pareto weight as a decreasing function of age, since younger agents dislike social security taxes.

With the arguments above in mind, I propose the following functional form for the Pareto weights:

$$\omega(\bar{e}_{zjt}, j) = \bar{e}_{zjt}^{\kappa_1} j^{\kappa_2}, \tag{1.14}$$

where $\kappa_1, \kappa_2 \in \mathcal{R}$ are parameters to be calibrated.

Of course, the specified Pareto weight function is ad-hoc. However, it captures important empirical observations. From a positive perspective, the social welfare function in eq. (1.13) is equivalent to a micro-founded probabilistic voting environment introduced by Lindbeck and Weibull, 1987.[20] The weight attached by a candidate to each group of voters reflects the propensity of voters to participate in political activities (voting, contributing money to elections, etc.). This weight corresponds to the Pareto weight ω. The specification of the Pareto weight function in eq. (1.14) reflects the data fact that participation in almost any form of political activities differs across household' age and income (see Bartels, 2009, Benabou (2000), Rosenstone and Hansen (1993)).

The calibrated values for κ_1 and κ_2 have no direct economic interpretation. This is why I construct the implied political Lorenz curve. It describes the proportion of political power held by the poorest fraction of population sorted by market income. The Lorenz curve implied by the model is plotted in figure 1.9. The 45 degree line corresponds to the case of equal Pareto weights.

---

[20]In this environment, there are two candidates, who are maximizing the probability of winning. Voters differ not only in terms of the most preferred social security policy but also in terms of their propensity to participate in political activities (vote, contribute money to elections, etc.). In equilibrium, both candidates propose the same policy, which maximizes the weighted sum of welfare of all voters.

To contrast political inequality in the model and in the data, I use two proxies for political inequality. The first proxy are voter turnout rates. It can be seen from the figure that the degree of political inequality implied by differential voter turnout rates is too weak as compared to the degree of inequality implied by the model. This is not surprising, since political power in the real world stretches far beyond pure participation in elections. The second proxy are campaign contributions. Figure 1.9 suggests that the model implied political inequality comes very close to the political inequality implied by an unequal distribution of campaign contributions.[21]

1.5   Quantitative exercise

I assume that in 1978 the economy is affected by the changes in the labor productivity process, the demographics, the income tax and transfer systems, and the Medicare. All these changes are consistent with the data. I update the corresponding parameters of the model assuming that all the remaining parameters, including the Pareto weights and the parameters of the pension benefit formula, stay at the level of the benchmark model economy. Table 1.4 summarizes the changes in the economic environment. Below I discuss them in more detail.

Labor productivity process

To account for the rise in cross-sectional earnings inequality, I make four modifications. First, I account for the widening college premium by updating the age-efficiency profiles of college and high school graduates, ζ_j, consistent with the CPS data during 1978-2005. Second, I account for an increasing share of college degree graduates, λ_H, from 19 percent in the benchmark model to 33 percent in 2005 (CPS). Third, I update the variance of the transitory component of the idiosyncratic labor productivity process, σ_v^2. Based on the time estimates by Heathcote et al. (2010b), σ_v^2 increased from 0.0560 in 1977 to 0.0735 in 2000. I assume that σ_v^2 increases linearly during this period and afterwards stays at the level of 2000 until the economy reaches the new steady state. Fourth, I calibrate the variance of the persistent component, σ_γ^2, to match the $P50/P20$ ratio of the distribution of lifetime earnings in the data in 2006. I assume that σ_γ^2 increases linearly during 1977-2006 and then stays at the level of 2006 until the economy reaches the new steady state.

---

[21]To construct figure 1.9, I use the Census data on voting in federal elections for the year 1980 from Table 13. *Reported voting and registration of primary family members by race, Spanish origin, age, and family income* available at https://www.census.gov/hhes/www/socdemo/voting/publications/p20/1980/ tab13.pdf. Households are sorted by market income in the Census data; this is why I also sort agents by income in the model. For campaign contributions, I borrow the computations obtained by (Bachmann and Bai, 2013) (2013, Table 1, Scenario II.), who used the *Cooperative Congressional Election Survey* from the years 2006 and 2008 (no data are available for 1980).

	Benchmark	Simulation	Source
1. Labor productivity process			
college premium, ζ_{jH}/ζ_{jL}	vectors	vectors	CPS
college share, λ_H	22%	32%	CPS
temporary shock, σ_v^2	0.056	0.0735	Heathcote et al. (2010b)
persistent shock, σ_γ^2	0.04	0.067	$P50/P20$ of lifetime earnings
2. Demographics			
birth rates, n	1.8%	1.3%	World Bank
3. Income tax and transfer programs			
average tax, m_0	0.48	0.30	Gouveia and Strauss (1994)
cons. floor for workers, $\underline{c}_W$	$14,515	$10,589	UKCPR
4. Medicare			
cap^M	$101,639	none since 1994	SSA

Table 1.4: Parameters changed in the simulation

Demographics

Falling birth rates have resulted in a higher old-age dependency ratio (i.e. ratio of retired to the working age population). To account for this, I update the birth rates in the model: relative to the benchmark model with the birth rate, n, of 1.8 percent, which is the average birth rate over the period 1960-77, the birth rate in the new steady state is assumed to be 1.3 percent.[22]

Income tax and transfer systems

The progressiveness of the income tax schedule declined sharply during 80s, mostly due to adoption of the Economic and Recovery Tax Act of 1981. For example, the top marginal income tax rate was cut from 70 to 50 percent. Using the estimates by Gouveia and Strauss (1994), I account for a drop in the average tax rate, m_0, from 0.48 in 1977 to 0.30 in 1989.

The means tested transfers for working households have decreased drastically since 1979 (figure 1.10). While the consumption floor was $14,515 in 1980, it fell down to $10,589 in 2011. I feed into the model the entire path of the transfers during 1980-2011.

Medicare

Consistent with the data, I account for the fact that the earnings cap for the hospital insurance, cap_M, was eliminated in 1994.

---

[22]The data on birth rates is from the World Bank, see http://data.worldbank.org/indicator/SP.DYN.CBRT.IN/countries/US?display=default.
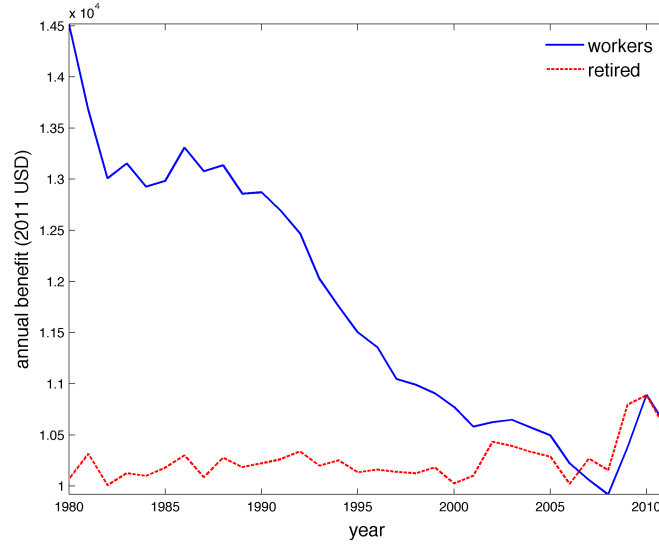
Figure 1.10: Dynamics of means-tested transfers for working and retired households in the data

1.6 Results

1.6.1 Benchmark model economy in 1977 and 2015 (before reform)

Figure 1.11 shows the dynamics of wages, total labor supply (in efficiency units), total pension benefits, interest rates, the social security tax rate and total lump-sum transfers, as the economy transits from the initial steady state to a new steady state under the changed parameters of the economic environment and the pension benefit formula as of 1977. The paths for the interest rate and the tax rate are given in percentage points, while the paths for other variables – in deviations from the initial steady state. The dashed vertical lines correspond to the year 2015.

As can be seen from the figure, the total supply of labor increases by more than 20 percent by 2015 due to the increased share of college graduates and the upward shift in the college premium. The total amount of pension benefits is ca. 7 percent above the level of the initial benchmark model economy. One reason for that is the rising dependency ratio due to falling fertility rates. It turns out that the surplus in social security contributions due to the increase in the total labor supply is not sufficient to cover the pension entitlements, so that the social security tax rate has to rise to almost 8 percent in order to balance the government budget.

The evolution of inequality in cross-sectional earnings in the model is shown in figure 1.12. I compute two measures of inequality: the $P50/P20$ ratio to measure inequality at the lower tail and the $P80/P50$ ratio to measure inequality at the upper tail of the pre-
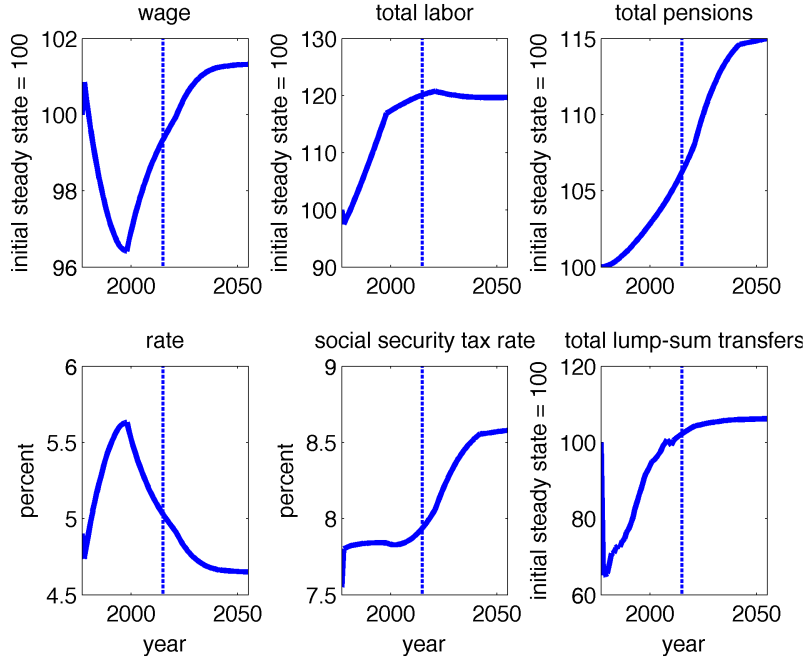
Figure 1.11: Transitional dynamics of aggregate variables in the model

government earnings distribution. From the figure we see that the cross-sectional earnings inequality increased both at the upper and lower tail of the earnings distribution. The increase at the top of the distribution was more pronounced (from ca. 2.10 in 1977 to ca. 2.75 in 2000) than at the bottom (from ca. 2.0 to ca. 2.35).

The transitional dynamics of lifetime earnings inequality in the model and in the data are fairly close as shown in figure 1.13. Consistent with the calibrated benchmark model economy, the inequality is measured among the individuals of (real life) age 46. The model matches almost exactly the levels of lifetime earnings inequality in 1977 and 2006 because they were the target moments for the calibration of the variances of the persistent shock, σ_γ^2. The model, however, generates a surprisingly realistic slope of the increase in lifetime earnings inequality, which was not targeted.

1.6.2   Optimal design of social security in 2015

I assume that in 2015 the government chooses the parameters of the pension benefit formula, (α_1, α_2). More precisely, the government solves the maximization problem in eq. (1.13), where the cumulative population density function is given by F_{2015}. The government chooses the policy under rational expectations about how this policy will affect the future equilibrium outcomes and the welfare of each agent who is alive as of 2015. Importantly, the parameters κ_1 and κ_2 of the Pareto weight function in eq. (1.14) are assumed to stay constant at the 1977 level (table 1.2).

The solution to the government's optimization problem is presented in table 1.5.
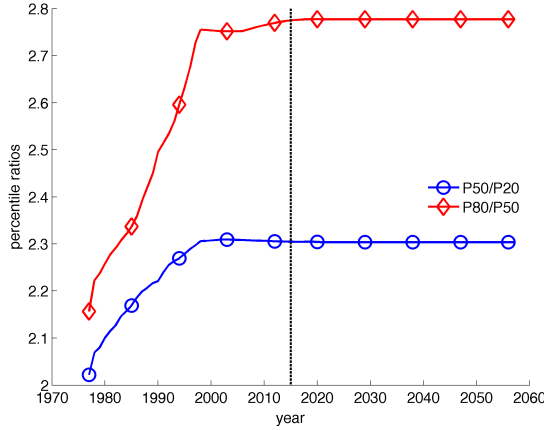
Figure 1.12: Transitional dynamics of cross-sectional earnings inequality in the model
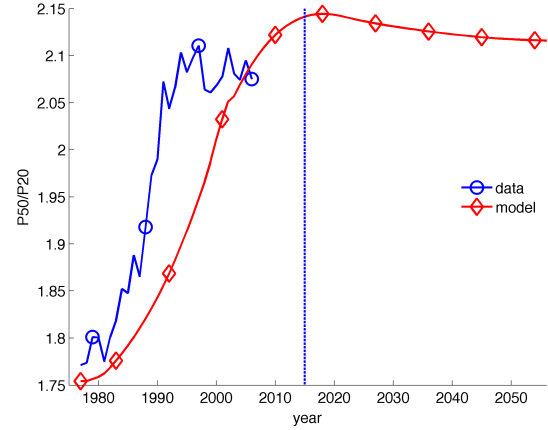
Figure 1.13: Transitional dynamics of lifetime earnings inequality in the model and in the data

	τ_t	α_1	α_2
1977	7.55%	0.514	0.577
2015 before reform	7.95%	0.514	0.577
2015 after reform	12.36%	0.77	0.40
2015 after reform (equal Pareto weights)	30.37%	1.79	−1.0

Table 1.5: Optimal design of social security in 2015

After the reform, the level of statutory replacement rates, α_1, goes up from 0.514 to 0.77, while the degree, to which pension benefits are linked to lifetime earnings, goes up because α_2 falls from 0.577 to 0.40.

Next I use the model to compute and compare the net transfers for newly born workers in 1977 and 2015 (before and after reform). As already pointed out, social security redistributes incomes over individual's lifetime. For a given worker, the lifetime redistributional impact depends on the difference between the present value of all contributions paid and the present value of all pension benefits received. To compute these net transfers, I first initialize a large panel of newly born workers and then simulate for each worker the paths of earnings, social security contributions, and pension benefits until the death of each individual, which is determined by the conditional survival probabilities. Then I compute each worker's net transfer in present value terms by discounting each payment in period t by the real interest rate, r_t. Finally, I sort the agents by their lifetime earnings and compute the average net transfer for each quintile of the lifetime earnings distribution. The results of this exercise are presented in table 1.6.

As can be seen from the table, the net transfers decreased significantly for all groups of

| | **1977** | **2015** | | **After reform** |
| | | Before reform | | |
	$	$	% change	$
$Q1$	59,580	54,176	−9.1	88,501
$Q2$	82,090	78,555	−4.3	116,943
$Q3$	84,050	77,835	−7.4	106,602
$Q4$	102,800	92,027	−10.5	119,794
$Q5$	100,960	88,450	−12.4	108,826

Table 1.6: Ex-post lifetime net transfers by quintiles of lifetime earnings distribution, 2011 U.S. dollars (discounted using r_t).

agents in 2015 (prior to the reform) as compared to the net transfers enjoyed by the agents born back in 1977. With negative 10.5 and 12.4 percent, the drop is most pronounced for the agents from the 4th and the $5th$ quintile of the lifetime earnings distribution, respectively.

1.6.3   Counterfactual analysis

I focus on two important changes: 1) the increased uncertainty in the labor productivity process (i.e. increased variance of the transitory and the permanent shocks), and 2) the upward shift in the college premium. I simulate the initial benchmark model economy by updating *only* the variances of the persistent and the temporary shocks in the first experiment, and the age-efficiency profiles of the high ability and low ability agents in the second experiment. All the remaining parameters are assumed to stay constant.

Figure 1.14 decomposes the dynamics of lifetime earnings inequality in each of the experiments.

1.7   Outlook

A lot of interesting and challenging questions remain to be explored in the future.

In this paper, I assume that each agent in the model economy corresponds to a *household* in the data. In the U.S., the pension benefits vary a lot by the *family* structure, since social security pays benefits not only to retired workers but also to their spouses, dependents and survivors. For example, a spouse can claim up to 50 percent of spouse's social security benefits, even if the spouse has never worked. Pronounced changes have occurred in family patterns since 1977: divorce rates have increased, while marriage rates have fallen. Besides, I focus on households composed of *male* heads only. It is well-known, though, that labor force participation rates among women have increased significantly since 1970s. Hence, fewer wives receive pensions as a function of their husbands' lifetime
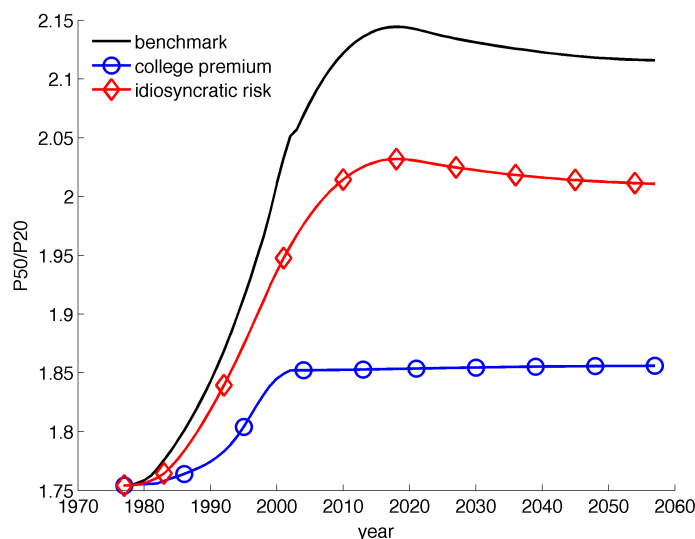
Figure 1.14: Transitional dynamics of lifetime earnings inequality in the benchmark model and counterfactuals

earnings but rather their own earnings histories. As suggested by Kopczuk et al. (2010), including females would reduce the overall inequality in average lifetime earnings and – in the context of my model – dampen the incentives of the government for more income redistribution. It is therefore important to understand how these trends in family structure and labor force participation affect the redistributive design of the social security system.

Furthermore, I only partially endogenize the social security system. For example, the maximum taxable earnings threshold is assumed to be constant. The wage inequality has increased dramatically at the upper tail of the (male) wage distribution since 1970s. So if the government has to increase the overall size of the pension system, it might adjust upwards the earnings cap rather than increase the social security tax rate.

Finally, it is important to recognize that social security is only one of several instruments to redistribute incomes. The redistributive design of the programs, such as Medicaid, Medicare, Supplemental Security Income, Earned Income Tax Credit, etc. has been changing a lot since 1970s. These programs are assumed to be exogenous in my model. It is important to analyze the trade-offs faced by the government when it can choose the redistributive design of several programs at the same time.

Chapter 2

Income inequality and political inequality in the U.S.

> *In a political system where nearly every adult may vote but where knowledge, wealth, social position, access to officials, and other resources are unequally distributed, who actually governs?*
>
> ROBERT DAHL (1961)

2.1 Introduction

The *median voter theorem* has been widely applied to account for the levels of income redistribution that we observe in the data. According to this theorem, if all households vote over income taxes in a referendum and every household has *equal* voting rights, then the median most preferred tax rate will beat any other alternative in a pairwise vote. However, prominent macroeconomic models by Krusell and Rios-Rull (1999), Klein et al. (2008), Corbae et al. (2009) predict a way too high equilibrium level of income taxes for the U.S. economy, when the tax rate is chosen by the median voter. To account for this discrepancy, Benabou (2000) departs from the assumption of equal voting rights. Instead, he assumes that the political process is biased towards richer agents. In this case, the decisive agent is richer than the standard median voter and therefore prefers a lower level of income redistribution, which is more consistent with the data.

    I extend a standard heterogeneous agents model with incomplete financial markets by introducing *wealth-weighted* voting over income redistribution. Political inequality is modeled as a simple functional relationship between private wealth holdings and agent's voting share. I set up a benchmark model economy, which matches the observed levels of earnings inequality in the U.S. in 1979. In particular, the wage process is calibrated to

match relatively low dispersion of wages and relatively low mobility across wage groups in the PSID data at that time. I use the model to identify the degree of political inequality, which is consistent with the level of income redistribution observed in late 1970s in the U.S. This allows me to uniquely pin down the state, in which the *weighted* median voter is located in the model in 1979. Then I conduct the following experiment. I reparameterize the wage process to match the increased cross-sectional earnings inequality in 1996 using PSID. Afterwards, I let agents vote in 1979 once-and-for-all on the degree of the income tax progressivity. This experiment imitates the adoption of the Economic and Recovery Tax Act of 1981, which significantly reduced the progressiveness of the statutory income tax schedule. Before agents vote, they anticipate the occurred changes in the wage process. I simulate the counterfactual time series of income redistribution since 1979 under the standard and the weighted median voters and compare them with the data.

I use a heterogeneous agents model with incomplete financial markets along the lines of Aiyagari (1994) and Huggett (1993). This modeling framework is reach in generating substantial heterogeneity in demand for income redistribution among agents. Moreover, it incorporates the demand for insurance as an important real-life feature. Demand for insurance becomes especially important when agents face higher income risk, which was the case in the data during1980s. As Quadrini (2009) puts it, "the use of a model with idiosyncratic risks provides a more solid and revealing answer to the question".

As a result of top coding in PSID, a model based on these data would not match the upper tail of the wealth distribution and therefore underestimate the wealth inequality. Since wealthier agents gain disproportionately more political power in the context of this model, it is important to achieve a good fit of the upper tail of the wealth distribution. This is why one important feature of the model is the presence of super-rich agents, whose productivity is calibrated to match the share of wealth held by the agents in the top quintile of the wealth distribution.

Throughout the paper, I measure total income redistribution as a difference between the market (pre-government) income inequality and the disposable (after-tax and after-transfer) income inequality. To provide a careful measure of income inequality, I use the $P50/P20$ percentile ratio to measure inequality at the bottom and the $P95/P50$ percentile ratio to measure the inequality at the top of the income distribution. I further decompose total income redistribution into the redistributive effect of *transfers* and the redistributive effect of *taxes*. The redistributive effect of transfers is given by the difference between market income inequality and inequality in disposable income without government taxes. Instead, the difference between the disposable income inequality and the inequality in disposable income without government taxes gives the redistributive effect of taxes.

Using the Current Population Survey, I show that total income redistribution reduced significantly both at the bottom and top of the income distribution during 1980s. Fur-

thermore, I show that *taxes* were responsible for the reduction in income redistribution and the bottom, while both taxes and transfers were driving the reduction at the top. This is why another important modeling feature are *progressive* income taxes, since proportional income taxes (used in the models mentioned above) would necessarily imply the redistributive effect of taxes equal to zero. In order to capture the progressiveness of the income tax system, I introduce the effective tax function by Gouveia and Strauss (1994). The progressivity parameter of this income tax schedule is endogenously determined in my model through voting.

There are several findings. First, I find a very high degree of political inequality in 1979. Whereas the standard median voter is located in the 2nd quintile of the equilibrium market income and earnings distributions of 1979, the weighted median voter is in the top decile of the market income distribution and 8th decile of the earnings distribution. I compare the political inequality in the model with the political inequality in the data, which I measure using voter turnout rates and campaign contributions. I find that the model implied inequality is stronger than the data implied inequality.

Second, the model with the identified weighted median voter is able to account for roughly 36 percent of the decline in income tax progressivity due to the Economic and Recovery Tax Act of 1981. Moreover, I show that in response to the increased earnings inequality a model, in which the standard median voter is decisive, predicts twice as much redistribution at the bottom and more two and a half times more redistribution at the top than in the data. On the contrary, a model, in which the weighted median voter is decisive, generates the redistributive effects of taxes and transfers, whose magnitudes are in line with the data.

Third, in order to understand the differences between the voting outcomes under the standard and the weighted median voters, I quantify the effect of rising dispersion of wages and rising mobility across wage groups since 1980 on the most preferred tax rate of each type of the voters. I show that agent's expectations about the future wage inequality are crucial in understanding the differences between the two equilibrium outcomes.

I proceed as follows. Section 2.2 emphasizes the contribution of this paper relative to the existing literature. Important data facts on income redistribution are shown in section 2.3. In section 2.4 I set up the model and define a politico-economic recursive competitive equilibrium. The parameters of the model are estimated in section 2.5. The quantitative exercise is done in section 2.6. Challenging tasks for future research are discussed in section 2.8.

2.2   Literature review

The current paper contributes to a vast strand of literature which attempts to explain the economic size of the government. In their seminal paper, Meltzer and Richard (1981) apply the standard median voter theorem to study the relationship between rising inequality and the resulting redistribution. Krusell and Rios-Rull (1999) show that the model by Meltzer and Richard (1981) considerably overestimates the amount of redistribution in the U.S. This is because the only distortionary effect of taxation comes from the endogenously supplied labor. Krusell and Rios-Rull (1999) extend this framework to a dynamic setup and add endogenous savings decisions. When agents vote, they internalize the distortionary effect of income taxation on the savings decision. Thanks to a stronger distortionary effect of taxation, the model explains better the amount of redistribution in the U.S. in 1995 but still overestimates it.

An important feature of the model by Krusell and Rios-Rull (1999) is that agents face no uncertainty and therefore agents have permanent differences in individual productivities. Benabou and Ok (2001) are the first to formalize an additional effect, which determines the voter's preferences over redistribution: the mobility across the earnings groups. The authors introduce idiosyncratic shocks to income and show that the longer the duration of the proposed tax reform, the less support for redistribution there will be. This is because the *prospect of upper mobility* makes some of the poor voters prefer low income taxation. However, the authors don't show empirically if mobility helps better understand the observed dynamics of redistribution in the U.S.

Corbae et al. (2009) are the first to do this. They introduce idiosyncratic productivity risk into the model by Krusell and Rios-Rull (1999). The authors ask what impact the increased mobility and the increased dispersion of earnings since 1980s in the U.S. had on redistribution. The authors find that all considered models, including the model with commitment and once-and-for-all voting by a median voter, overestimate the amount of redistribution. In this paper I show that a model, which accounts for the wealth bias, has considerably stronger predictive power.

In order to explain the *redistribution puzzle*, Benabou (2000) models the idea of a wealth bias in the political system. He associates the wealth bias with a wealth-weighted voting by introducing a simple functional relationship between private wealth holdings and the vote share. He shows analytically that high inequality can be consistent with low redistribution. High inequality increases redistribution, as in the standard model, but redistribution may in turn reduce inequality. Benabou (2000) explores the possibility of multiple equilibria raised by the interaction of these two avenues of causality. Even though very elegant, his model is too stylized to be calibrated to the U.S. data.

Lagunoff and Bai (2013) adopt a *revealed preference* approach to the question of what

can be inferred about the bias in a political system when an outside observer only observers a sequence of taxation policies but doesn't observe the underlying distribution of political power. They assume the same functional form of the bias as Benabou (2000). They establish a *Universal Bias Principle* that says that if agent's indirect utility function is single-peaked in the tax rate, *any* level of the wealth bias can rationalize *any* observed level of taxation. To ensure uniqueness of the wealth bias, information on agent's preferences from the polling data is needed. I generate such information from the model and pin down uniquely the degree of political inequality in the U.S.

Bachmann and Bai (2013) are the first to introduce the wealth bias into a structural model, which they then estimate. They show that the observed contemporaneous correlation between output and government purchases in the U.S. can be better explained in a model, which accounts for the wealth bias in the political process. In their model the government spending is utility enhancing, and wealthier individuals desire higher levels of government spending. During economic booms agents with lower wealth gain political power, which dampens the co-movement of government spending and the output. Since the authors abstract from any redistribution, their model cannot be used to study the impact of the wealth bias on after-tax income inequality. Besides, the authors consider a government, which chooses the amount of government spending so as to maximize a weighted social welfare function, with weights dependent on the wealth of the households. I do not use this type of social choice institutions, since it leads to a bad fit in the data as shown by Corbae et al. (2009). Instead, I assume majoritarian voting.

2.3   Data facts: Income redistribution

Income redistribution in the U.S. takes many forms, such as progressive income taxes, means-tested welfare programs, etc. One way to measure the *overall redistributive effect* of taxes and government transfers is to compare the market (pre-government) income inequality with the disposable (after-tax and after-transfer) income inequality. The market income includes earnings, business and farm income, capital interest, dividends and private transfers, while the disposable income adds to the market income the government transfers (unemployment insurance and welfare) and subtracts tax liabilities. Thus, the larger the difference between the market income and the disposable income, the more effective the taxes and the transfers are in reducing income inequality.

The overall redistributive effect can be further decomposed into the redistributive effect of transfers and the redistributive effect of taxes. The *redistributive effect of transfers* is given by the difference between market income inequality and inequality in disposable income without government taxes. Instead, the difference between the disposable income inequality and the inequality in disposable income without government taxes gives the

redistributive effect of taxes.

There are also many ways to measure income inequality. I use the $P50/P20$ percentile ratio to measure inequality at the bottom and the $P95/P50$ percentile ratio to measure the inequality at the top of the income distribution.

Figures 2.1 and 2.2 plot the dynamics of inequality in market income, disposable income and market income after transfers for the bottom and the top of the income distribution, respectively.[1] Several observations are worth noting. First, inequality in market income was increasing steeply both at the bottom and at the top of the income distribution during 1980s. Second, transfers are responsible for the most of the reduction in income inequality at the bottom. Conversely, most income redistribution at the top is due to taxes because of their progressive nature. With a proportional income tax system the redistributive effect of taxes would be zero. Third, the inequality in disposable income both at the bottom and at the top mirrors the market income inequality during 1980s.
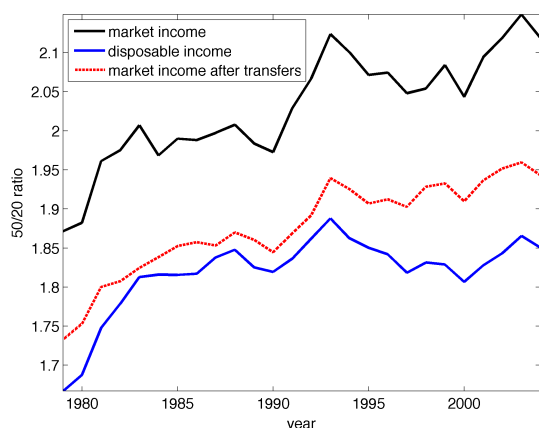
Figure 2.1: Dynamics of inequality at the bottom of the income distribution
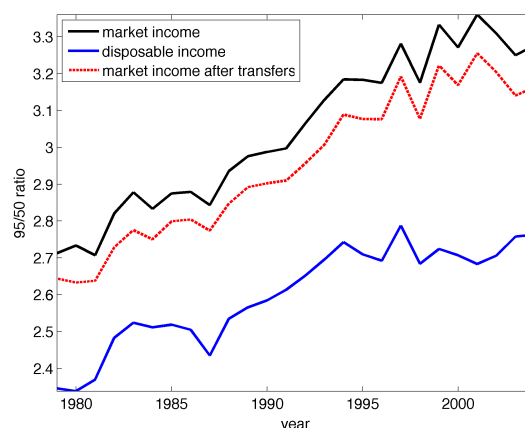
Figure 2.2: Dynamics of inequality at the top of the income distribution

Using the time series in the previous two figures, I further construct figure 2.3 to show explicitly the redistributive effects of taxes and transfers during 1980s. I normalize the redistributive effects to be 1.0 in 1980. For example, the redistributive effect of taxes of 0.2 at the bottom of the income distribution in 1983 means that taxes reduced the gap between the disposable income and the disposable income without government transfers by 80 percent less than in 1980.

---

[1] The time series for income are taken from the Current Population Survey. I restrict attention to households with a male head of working age (20-64), who works at least 260 hours per year. Household's income is adjusted using the OECD equivalence scale. Transfers include income from public assistance or welfare, unemployment insurance, social security or railroad retirement, supplemental security, and educational assistance. Tax liabilities include federal and state income tax liabilities, social security payments and earned income tax credit. Tax liabilities are computed using TAXSIM.
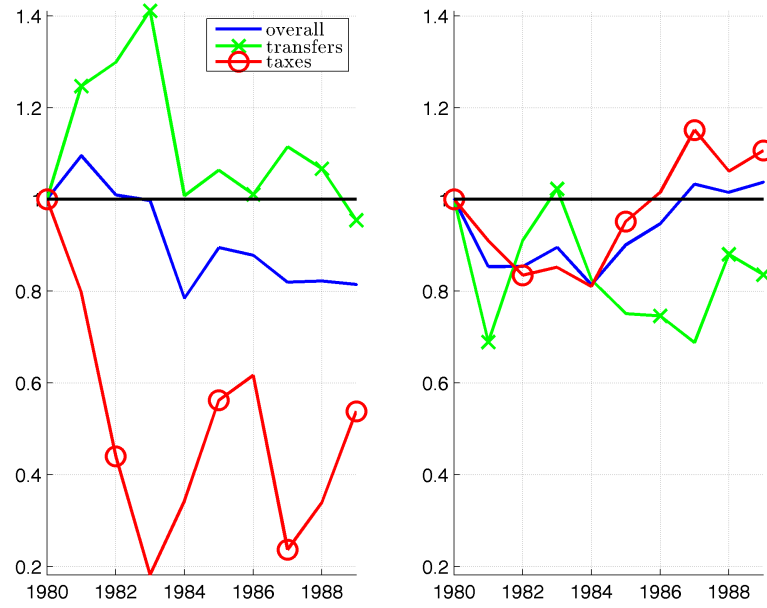
Figure 2.3: Dynamics of redistribution at the bottom (left) and the top (right) of the income distribution

2.4   The model

2.4.1   Endowments and preferences

There is a unit measure of infinitely lived households. In each period households are endowed with one unit of productive time. They spend this time supplying labor to a competitive labor market or consuming leisure. Households are ex-ante identical. The only source of ex-post heterogeneity is exogenous idiosyncratic labor productivity shocks. At any date t agent's productivity, denoted by $\epsilon \in E = \{\epsilon_1, \ldots, \epsilon_z\}$, follows a first-order Markov process with a $z \times z$ transition matrix $\Pi(\epsilon' \mid \epsilon)$, which gives the probability of moving from the state ϵ to the state ϵ'. All agents face the same stochastic process for productivity. Agents know the law of motion for the productivity process.

There are no insurance markets against the idiosyncratic risk. In order to self-insure against the risk of low labor productivity in the future, agents can save in capital. Individual capital holdings are denoted a.[2] Self-insurance is limited by the assumption that the households can not borrow, i.e. $a \geq 0$.[3] Therefore, the individual state space consists of the sets $(a, \epsilon) \in [0, \infty) \times E$. Let $F_t(a, \epsilon)$ denote the cumulative density function of the assets and productivities in the economy at date t. The corresponding density function is denoted by $f_t(a, \epsilon)$.

---

[2]Since I consider only one type of asset, I refer to a as capital, wealth and assets interchangeably.

[3]This is a common assumption in the related literature, see Domeij and Heathcote (2004), Conesa and Krueger (2006).

The instantaneous utility function of the agent reads:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \tag{2.1}$$

where c is agent's consumption and σ controls the degree of relative risk aversion.

2.4.2 Technology

The aggregate technology is represented by a standard Cobb-Douglas production function $K_t^\theta N_t^{1-\theta}$, where K_t, N_t denote the aggregate capital stock and the aggregate efficient labor input, respectively, and θ is the capital share in production. The capital stock depreciates at the rate δ. As standard with a constant returns to scale technology and perfect competition, I assume the existence of a representative firm operating this technology.

2.4.3 Government policy

Define

$$y_t(a, \epsilon) = \epsilon w_t + r_t a_t, \tag{2.2}$$

the pre-tax income of the agent in the state (a, ϵ) at time t, where r_t is the risk-free interest rate and w_t is the wage per efficiency unit of labor.

In every period the government taxes each agent's market income. Among a large set of different specifications for the taxation rule, I choose the progressive taxation rule by Gouveia and Strauss (1994). The authors derived the rule from the theory and estimated its parameters for the U.S. for the period 1979-89. This progressive taxation rule reads:

$$M(y; \tau) = \tau[y - (y^{-m_1} + m_2)^{-1/m_1}], \tag{2.3}$$

where $M(y)$ is the amount of taxes the agent has to pay if her pre-tax income equals y. The marginal tax rate is given by $M'(y)$ and the average tax rate is given by $M(y)/y$. Parameters τ and m_1 set the type and the magnitude of taxation, figures (2.4)-(2.5).

Parameter $\tau \in [0, \bar{\tau}]$ is the marginal (and average) tax rate as income goes to infinity, i.e. $\tau = \lim_{y \to \infty} M'(y) = \lim_{y \to \infty} M(y)/y$. Parameter m_1 determines the curvature of the marginal tax function $M'(y)$. This parameter has a meaningful economic interpretation on the whole set $[-1, \infty)$: if $m_1 = -1$, taxes are independent of income and equal τm_2; with $m_1 \to 0$ one obtains a proportional taxation rule $M(y) = \tau y$; for $m_1 > 0$, there is a progressive taxation of incomes. Parameter m_2 is a scaling parameter.

I choose the taxation rule by Gouveia and Strauss (1994) because it defines an empirically realistic progressive taxation schedule, and as I showed in section 2.3 the redistribu-
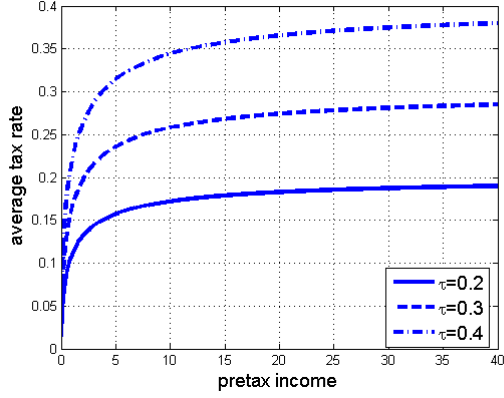
Figure 2.4: Average tax rates for different τ (with $m_1 = 0.768$ and $m_2 = 0.55$)
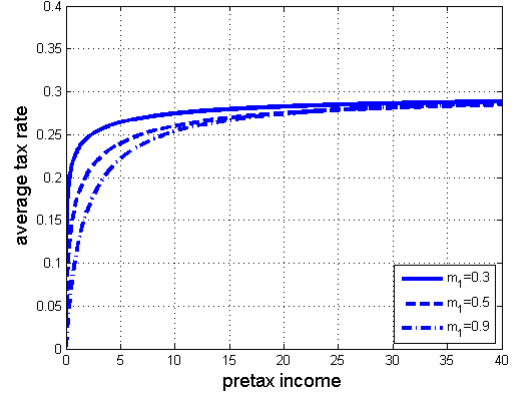
Figure 2.5: Average tax rates for different m_1 (with $\tau = 0.3$ and $m_2 = 0.55$)

tive effect of taxes is quantitatively important for the U.S. economy, especially during the 1980s.

In every period, the government provides lump-sum transfers, T_t, and spends G_t on expenditures, which are wasted in the context of the current model. The government runs a balanced budget, so that the following equality holds for all t:

$$T_t + G_t = \sum_{\epsilon} \int M(y_t(a, \epsilon); \tau_t) f_t(a, \epsilon) \mathrm{d}a. \tag{2.4}$$

2.4.4 Recursive competitive equilibrium

I define the politico-economic recursive competitive equilibrium in two steps. In the first step, I set up a recursive competitive equilibrium, when agents take the tax rate τ as given. In the second step, I make τ consistent with the political process. The time index t is dropped for the ease of notation.

Definition. Given a rate of taxation τ, a **recursive competitive equilibrium** is a set of functions $\{v, c, a', T, G, w, r, F\}$, such that:

- Given (τ, w, r, T, c, a'), v solves the Bellman equation:

$$v(a, \epsilon, F; \tau) = \max_{c, a'} \left\{ u(c) + \beta \sum_{\epsilon'} \Pi(\epsilon' \mid \epsilon) v(a', \epsilon', F'; \tau) \right\}$$

  subject to:

$$a' = [\epsilon w(K, N) + (1 + r(K, N))a] + T - M(y(a, \epsilon); \tau) - c,$$
$$a', c \geq 0, \tag{2.5}$$

  where $u(c)$, y and $M(\cdot)$ are defined in (2.1), (2.2) and (2.3), respectively. The parameter β is the subjective discount factor. A prime denotes tomorrow's realization

of the related variable. The object $v(a, \epsilon, F; \tau)$ denotes the discounted life-time in-
direct utility of a household in the state (a, ϵ), when the aggregate distribution of
agents is F and the tax rate is τ.

- The pricing functions satisfy: $r = \theta K^{\theta-1} N^{1-\theta} - \delta$ and $w = (1-\theta) K^\theta N^{-\theta}$.

- The distribution of agents across productivity levels and assets evolves according to

$$F'(\epsilon', a') = \sum_\epsilon \Pi(\epsilon' \mid \epsilon) F(\epsilon, (a(\epsilon, a'))^{-1}),$$

  where $(a(\epsilon, a'))^{-1}$ is the inverse of the optimal policy $a' = a'(\epsilon, a)$ with regard to
  current period wealth a.

- The government runs a balanced budget according to (2.4).

- The markets clear: $N = \sum_\epsilon \int \epsilon f(a, \epsilon) \mathrm{d}a$ (labor market), $K' = \sum_\epsilon \int a'(a, \epsilon) f(a, \epsilon) \mathrm{d}a$
  (asset market), $K^\theta N^{1-\theta} = \sum_\epsilon \int c(a, \epsilon) f(a, \epsilon) \mathrm{d}a + K' - (1-\delta) K$ (goods market).

Definition. A **steady state recursive competitive equilibrium** is a recursive com-
petitive equilibrium with $F' = F$.[4]

2.4.5   Politico-economic recursive competitive equilibrium

I assume a particular form of the social choice institutions – a majority voting system.[5]
Agents vote on the design of the progressive taxation schedule specified in (2.3). I assume
that agents vote on the parameter τ because it affects the progressivity of the taxation
schedule (figure 2.4), which is the key variable in the paper. The second reason is that I
am able to show that the key assumption of the median voter theorem (single peakedness
of the indirect utility function) is satisfied if the voting is on the parameter τ.

I assume that the government can fully commit to keep the chosen tax rate constant.
This is a restrictive assumption but relaxing it complicates the computational procedure
considerably. Since this paper is the first attempt to explain the dynamics of redistribution
when political power is unequally redistributed, it seems appropriate to start with a
simpler case of full commitment.

I assume that the weight attached to the agent's vote depends on agent's capital
holdings. More specifically, the vote allocation to the single agent in the state (a, ϵ) is
given by a^α, with $\alpha > 0$ governing the strength of the wealth bias. Then the weight

---

[4]Huggett (1993) (pp. 960-961) provides conditions for the existence and the uniqueness of the station-
ary equilibrium in a similar model. He also describes the procedure to find the stationary equilibrium on
a computer.

[5]I choose the majority voting system, in order to relate my paper to the existing papers on the size of
the government mentioned in section 2.2.

attached to all agents in the state (a, ϵ) is the product of the mass of agents in this state and the weight attached to each agent in this state, $f(a, \epsilon) \times a^{\alpha}$.

There are at least three explanations for why richer people may have stronger power in politics than the poor. The first explanation is that rich people tend to vote more often than the poor. Rosenstone and Hansen (1993) and Page et al. (2013) find that propensity to participate in voting in the U.S. rises with income and wealth. The second explanation is that rich people are able to make higher campaign contributions and potentially affect the election outcome or to obtain influence over legislative decision-making by the successful candidate. Rosenstone and Hansen (1993) and Campante (2011) find that propensity to make contributions rises with income. The third explanation is that rich people tend to meet public officials more often than the poor. About half of respondents in Page et al. (2013) reported contacting at least one highly ranked official within a half year, a much higher proportion than among the general public.[6]

I find the most preferred tax rate of agents in each state as follows: among all possible $\tau \in [0, \bar{\tau}]$, the agent in the state (a, ϵ) prefers the tax rate $\hat{\tau}$, which maximizes her ex ante life-time utility given by $v(a, \epsilon, F; \hat{\tau})$. Then I order the obtained tax rates $\hat{\tau}$. The median voter is then the agent type, whose most preferred tax rate is the weighted median of the ranked most preferred tax rates. Below I refer to this agent type as the **weighted median voter** and denote her most preferred tax rate $\tau^{\star}$.

Definition. The **steady-state politico-economic recursive competitive equilibrium** is:

- A set of functions $\{c, a', T, G, w, r, F, v\}$ that, given τ, satisfies the definition of the steady-state recursive competitive equilibrium above.

- The bias $\alpha^{\star}$, such that the policy outcome is $\tau = \tau^{\star}$, where $\tau^{\star}$ satisfies:

$$\frac{\int \mathcal{I}_{\{(a,\epsilon):\hat{\tau}(a,\epsilon) \geq \tau^{\star}\}} a^{\alpha^{\star}} f(a, \epsilon) \mathrm{d}a}{\int a^{\alpha^{\star}} f(a, \epsilon) \mathrm{d}a} \geq \frac{1}{2} \text{ and } \frac{\int \mathcal{I}_{\{(a,\epsilon):\hat{\tau}(a,\epsilon) \leq \tau^{\star}\}} a^{\alpha^{\star}} f(a, \epsilon)}{\int a^{\alpha^{\star}} f(a, \epsilon) \mathrm{d}a} \geq \frac{1}{2}, \quad (2.6)$$

with $\mathcal{I}$ – an indicator function[7] and $\hat{\tau}(a, \epsilon)$ given by:

$$\hat{\tau}(a, \epsilon) = \arg \max_{\tau \in [0, \bar{\tau}]} v(a, \epsilon, F; \tau). \quad (2.7)$$

---

[6]The study by Rosenstone and Hansen (1993) is based on National Election Studies in the U.S. in 1952-1988. Page et al. (2013) interview 83 top wealthy households in Chicago area and find that 99 percent of the respondents participate in voting. Campante (2011) uses the data on contributions given to presidential candidate committees and to party committees from the Federal Election Commission and the 2006 Cooperative Congressional Survey. He finds in a probit regression that individuals with a higher income have a higher probability to make campaign contributions and that, conditional on contributing, income increases the value that is donated. Data show that fewer than 25 percent made contacts (in a broader sense) within a year.

[7]This condition just determines the weighted median of the ranked most preferred tax rates. It is convenient to re-scale each weight $a^{\alpha} f(a, \epsilon)$, so that all the weights sum up to one. This explains the term $\int a^{\alpha} f(a, \epsilon) \mathrm{d}a$ in the denominator.

The median voter exists if each agent's indirect utility $v(a, \epsilon, F; \tau)$ is single-peaked in τ. Single-peakedness ensures that the median ranked preferred tax rate beats any other tax rate in pairwise comparison. I do not have a proof of existence of the voting equilibrium. Neither am I aware of any existence results for the incomplete market literature. However, I verify numerically that in every simulation the indirect utility function satisfies single peakedness for every (a, ϵ).[8]

2.5 Calibration of the benchmark model

I assume that the model economy is in a steady state that matches key observations of the U.S. economy in 1979-80. The parameters of the model can be summarized in three different sets: (i) preferences and technology $\{\beta, \gamma, \sigma, \theta, \delta\}$; (ii) government parameters $\{\tau, \bar{\tau}, m_1, m_2, G\}$; and (iii) employment process $\{E, \Pi\}$. I take some of these parameter values from the data and the related studies, whereas some of the parameters are found by moment matching. The values for the parameters determined outside of the model are summarized in table 2.1, while the calibrated parameters are shown in table 2.2. The model period is one year.

2.5.1 Calibration with an exogenous policy

Preferences and technology

The discount factor $\beta = 0.954$ is chosen so that the equilibrium of the benchmark economy implies a capital-output ratio of 3.3. I fix the parameter σ, which controls the degree of relative risk aversion, to 1.0. I set the labor's share of income equal to 64 percent (which implies $\theta = 0.36$) and the rate of depreciation δ to 6 percent.

Government parameters

The political element of the model is not exploited at this stage, since I treat the rate of taxation τ (and the potential bias in the political process) as exogenously given.

I set G to match the share of government purchases and spending on social security (which is outside the model) of 14.3 percent in 1979.[9]

---

[8]Azzimonti et al. (2006) show analytically the existence of the median voter for a complete market model with a proportional income tax rate.

[9]The data come from Table 15.5. "Total Government Expenditures by Major Category of Expenditure as Percentages of GDP: 1948–2006" on the U.S. Government Printing Office web page under Budget of the United States Government: Historical Tables Fiscal Year 2008, see http://www.gpoaccess.gov/usbudget/fy08/sheets/hist15z5.xls.

Using the estimates by Gouveia and Strauss (1994), I set τ to 0.467 and the curvature of marginal taxes m_1 to 0.823, which are averages of the corresponding estimates for 1979-80. The parameter m_2 is calibrated to match the profile of effective income tax rates by the quintiles of the income distribution in 1979 (see below).

Employment process

The employment process is at the core of my analysis. I need to parameterize the transition matrix Π as well as the vector of employment shocks E. The starting point is the annual mobility matrix from 1978-79 estimated by Corbae et al. (2009) based on the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal data set particularly useful for measuring the evolution of earnings over time but earning are top coded. As a result of top coding, a model based on these data doesn't match the upper tail of the wealth distribution and therefore underestimates the wealth inequality.

There are two reasons why it is important to have an accurate measure of the top tail of the wealth distribution in my model. First, wealthier agents gain disproportionately more political power; therefore, existence of the very wealthy agents is important for the voting outcome. Second, below I have to resort to the data from the Survey of Consumer Finances (SCF) and the data by the Internal Revenue Service. As opposed to the PSID, these two data sources account sufficiently well for the extreme wealth inequality, so that combining these data with the PSID data might lead to wrong results.

In order to account for the extreme wealth inequality, I use an approach similar to Domeij and Heathcote (2004), Castaneda et al. (2003) and Diaz and Luengo-Prado (2010). I divide the households into two wage groups: super-rich and regular households. For regular households, the idiosyncratic productivities follow an autoregressive process $\log(\epsilon_{t+1}) = \rho \log(\epsilon_t) + u_{t+1}$. The shock u_{t+1} is iid with mean zero and variance given by $(1 - \rho^2)\bar{\sigma}^2$, where $\bar{\sigma}^2 = var(\log(\epsilon_{t+1}))$. In the discussion to follow, I associate the parameter $\bar{\sigma}^2$ with the dispersion of idiosyncratic shocks and the parameter ρ with the mobility across productivity levels with higher ρ meaning lower mobility.

Using the PSID data for 1979-80, I set the persistence coefficient ρ to 0.77 and the dispersion of productivities $\bar{\sigma}^2$ to 0.75. I approximate this autoregressive process by a 5-point Markov chain using the procedure by Tauchen (1986). This gives me a vector of discrete productivity levels $\{\epsilon_1, ..., \epsilon_5\}$ and a 5×5 transition matrix. The productivities ϵ are normalized, so that $\epsilon_3 = 1$. In the approximated process, the dispersion of productivities is reflected by the variance of the discrete values of ϵ, while the mobility across the productivity levels is reflected by the diagonal elements of the transition matrix with higher probabilities meaning lower mobility.

To account for the super-rich, I add to the obtained transition matrix the 6th row

Parameter	Description	Value	Source
Preferences:			
σ	degree of risk aversion	1.0	fixed
Technology:			
θ	capital share	0.36	Conesa and Krueger (2006)
δ	depreciation	6	Corbae et al. (2009)
Productivity process for regular households:			
ρ	persistence	0.77	PSID
$\bar{\sigma}$	variance	0.75	PSID
Government policy:			
τ	marginal tax for ∞ income	0.467	Gouveia and Strauss (1994)
m_1	curvature	0.823	Gouveia and Strauss (1994)

Table 2.1: Parameters determined outside the model

and the 6th column. I assume that each regular household has an equal probability of 0.0003 of becoming a super-rich, and the probability that a super-rich remains her status is 0.95, while she can draw ϵ_1 with probability 0.05. These values imply that in the long run the super-rich represent 0.6 percent of the total population. I add to the vector of productivities $\{\epsilon_1, ..., \epsilon_5\}$ the shock for the super-rich, ϵ_6, which I estimate to match the share of wealth held by the top quintile of the wealth distribution, which was 81.3 percent in 1983.[10]

The resulting 6×6 transition matrix Π together with the vector of productivities E are given in table 2.6 in Appendix. To quantify the degree of mobility, I compute the expected durations of each shock ϵ_n, which is a reciprocal of $1 - \Pi(\epsilon_n \mid \epsilon_n)$. The expected duration of being super-rich is 20 years.

Parameter	Description	Target	Value
β	discount factor	$K/\text{GDP}=3.3$	0.954
ϵ_7	productivity of super-rich	wealth held by $Q5 = 81.3\%$	$\epsilon_1 \times 1100$
G	wasted government spending	$G/\text{GDP} = 14.3\%$	0.87
m_2	scale factor	effective tax rates (table 2.4)	0.14
$\alpha^\star$	wealth bias	$\tau^\star = 0.467$	1.26

Table 2.2: Calibrated parameters

---

[10]Wealth is measured as net worth as defined in Wolff (2004). More specifically, net worth is the difference in value between total assets and total liabilities or debt. Total assets are composed of financial and non-financial assets. Financial assets do not include future social security benefits. Non-financial wealth includes (among others) housing. Total liabilities include (among others) mortgage debt. I choose net worth and not financial wealth because net worth reflects wealth as a store of value and therefore a source of potential consumption. Thus, net worth is of higher importance to households than financial wealth, when it comes to voting over the future permanent income tax rate (this will be the quantitative exercise below).

Moment	Benchmark	U.S.	Moment	Benchmark	U.S.
Wealth distribution:			**Market income distribution:**		
- Gini	0.76	0.79	- $P50/P20$	1.9	1.9
- Percentiles:			- $P95/P50$	3.3	2.7
$Q1$	0.6	0	**Disposable income distribution:**		
$Q2$	3.3	0.9	- $P50/P20$	1.5	1.7
$Q3$	6.7	5.2	- $P95/P50$	2.4	2.3
$Q4$	11.3	12.6	**Earnings distribution:**		
$Q5$	78.0	81.3	- $P50/P20$	1.9	1.8

Table 2.3: Comparing the calibrated benchmark model economy with the data in 1979

2.5.2   Model fit

Table 2.3 evaluates the performance of the calibrated benchmark model economy. First, the model does a very good job in matching inequality at the bottom of the market income distribution, while inequality at the top is overestimated. Inequality both at the top and the bottom of the disposable income distribution is fairly close to the data, which is surprising because it was not targeted. Furthermore, the benchmark model economy does a relatively good job of accounting for the shares of wealth owned by the households in the upper tail of the wealth distribution.

Table 2.4 compares the effective tax rates defined as the ratio of total taxes to the market income after transfers.[11] As it can be seen from the table, the model manages to achieve a sensible fit of the effective tax rates in the data.

Quintile	Model	Data
$Q1$	2.4	0.0
$Q2$	5.0	4.4
$Q3$	7.1	8.2
$Q4$	10.7	11.1
$Q5$	18.2	18.8

Table 2.4: Effective tax rates in the model and the data (1979), percent

2.5.3   Calibration with an endogenous policy

In this section, I find the strength of the wealth bias, $\alpha^\star$, which rationalizes the tax rate $\tau_B^\star = 0.4670$ observed in the data in 1979 (table 2.1), where B stays for *bias*. Below

---

[11] I use the data from "Historical Effective Tax Rates, 1979 to 2005" (Table 5. "Total Income and Total Federal Tax Liabilities for All Households, by Household Income Category, 1979 to 2005") by the Congressional Budget Office, which are available at http://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/98xx/doc9884/12-23-effectivetaxrates_letter.pdf. Market income is defined as a sum of wages, proprietors' and other business income, interest and dividends, and imputed taxes. I exclude capital gains from the definition of market income as they are not part of the model. I define transfers as cash transfers and in-kind benefits. For taxes, I take individual income taxes. All numbers are for 1979.
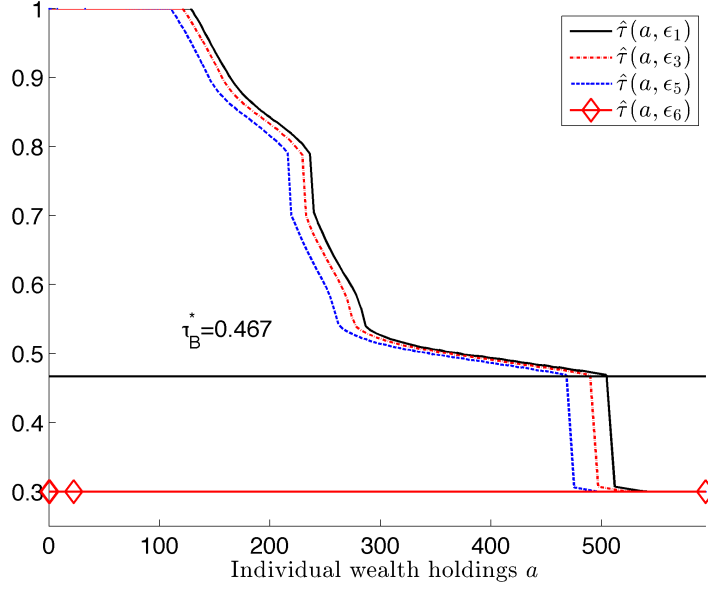
Figure 2.6: Most preferred tax rate $\hat{\tau}(a, \epsilon)$

I proceed as follows. I first explore agent's most preferred tax rates captured by the function $\hat{\tau}(a, \epsilon)$. Then I identify the weighted median voters in the model. I also identify the standard (unweighted) median voter and her most preferred tax rate, $\tau_U^\star$, where U stays for unweighted. The computational strategy for each of the steps is explained in Appendix.

Agent's most preferred tax rate

Agent's voting decision in 1979 can be better understood from figure 2.6, where I plot agent's most preferred tax rate as a function of agent's asset holdings and productivity (ignore for a moment the vertical and horizontal lines). Two observations are worth noting. First, for a given level of asset holdings, households facing a lower employment shock prefer a higher tax rate. Agents anticipate that the employment process in the future is going to be persistent, so that low productivity agents today expect to face low productivity in the future. Second, the preferred tax rate is a weakly decreasing function of private wealth. Note that the most preferred tax rate of the super-rich is $\hat{\tau}(a, \epsilon_7) = 0.3$ for all a.[12]

Revealed political power

This identification of the political bias α follows closely the reverse engineering approach discussed by Lagunoff and Bai (2013), who have the same specification for the wealth bias

---

[12]I restrict the government to sustain the G/Y ratio of the calibrated benchmark model economy (table 2.2). The lowest possible tax rate to satisfy this requirement is $\tau = 0.3$.

as in this paper. Their *Universal Bias Principle* ensures that if agent's indirect utility function is single-peaked in τ, then in principle *any* level of the wealth bias can rationalize *any* observed level of taxation $\tau^\star$. To guarantee the uniqueness of the wealth bias for a given equilibrium level of taxation, information on agent's preferences from the polling data is needed. In my model, the function of most preferred tax rates $\hat{\tau}(a, \epsilon)$ delivers such information.

The identified value of the parameter $\alpha^\star$ is shown in table 2.2. This value allows me to uniquely pin down the state, in which the weighted median voter is located in the model in 1979. To identify the standard median voter, I set $\alpha^\star$ to 0 in (2.6). It turns out the most preferred policy of the standard median voter is $\tau_U^\star = 1.0$.

Whereas the standard median voter is located in the 3rd quintile of the equilibrium wealth distribution of 1979, the weighted median voter is the agent in the top decile of the wealth distribution.[13] Furthermore, I find that the standard median voter is located in the 2nd quintile of the equilibrium market income and earnings distributions of 1979, while the weighted median voter is in the top decile of the income distribution and 8th decile of the earnings distribution.

The obtained value for the strength of the wealth bias $\alpha^\star$ has no direct economic interpretation. This is why I construct a political Lorenz curve implied by $\alpha^\star$, since it gives a simple measure of political inequality. More precisely, it describes the proportion of political power held by the poorest fraction of population sorted by market income.[14] The Lorenz curve is plotted as a solid line in figure 2.7. The line of complete political equality (the 45 degree solid line) corresponds to the case of one agent-one vote.

It is interesting to contrast the model results regarding political inequality with the data. In the data, though, we lack direct measures of political inequality. We could instead use some indirect measures. Recall from the introduction that there are different explanations for why rich people have stronger power in politics than the poor in the U.S. The first explanation is that rich people tend to vote more often than the poor. Thus, one could use participation rates in elections as a measure of political inequality. The resulting political Lorenz curve is represented in figure 2.7.[15]

It can be seen from the figure that the degree of political inequality implied by the participation rates is too weak as compared to the degree of inequality implied by the

---

[13]As can be seen from figure 2.6, apart from the weighted median voter, there are agents in other states whose most preferred tax rate is $\tau_B^\star$. However, no value of α is consistent with the fact that these agents are pivotal in equilibrium given the initial distribution of agents $f_0(a, \epsilon)$ and the function of most preferred tax rates $\hat{\tau}(a, \epsilon)$.

[14]I sort agents by market income in order to allow for a comparison with the data (see footnote 15).

[15]To construct this political Lorenz curve, I use the Census data on voting in federal elections for the year 1980 from Table 13 "Reported voting and registration of primary family members by race, Spanish origin, age, and family income" available at https://www.census.gov/hhes/www/socdemo/voting/publications/p20/1980/tab13.pdf. Households are sorted by market income in the Census data; this is why I also sort agents by income in the model.
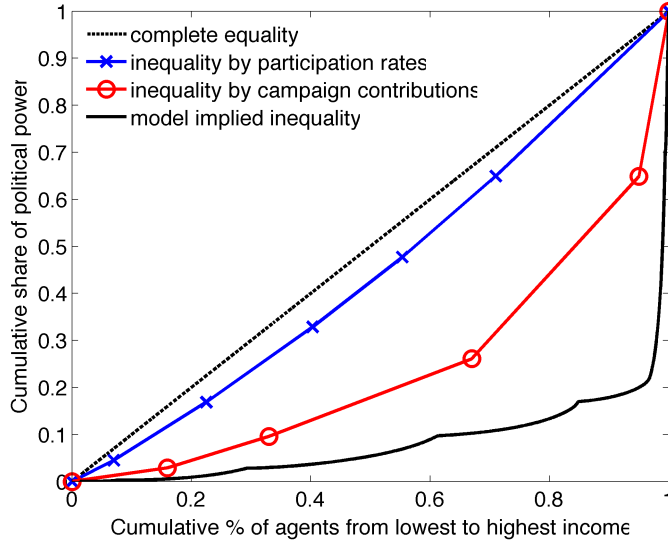
Figure 2.7: Political Lorenz curve

model. This is not surprising, since the political power in the real world stretches far beyond pure participation in elections. Recall from the introduction that the second explanation for why rich people have stronger political power is that rich people are able to make higher campaign contributions and potentially affect the election outcome or to obtain influence over legislative decision-making by the successful candidate. Following Bachmann and Bai (2013), I assume that one dollar of campaign contributions buys one vote and that the mapping between agent's wealth a and her campaign contributions (and therefore, de facto votes) is given by the function a^{α}. Then I use the data on campaign contributions to compute the political Lorenz curve, which is represented by a dot-dashed line in figure 2.7.[16]

Figure 2.7 suggests that de facto political inequality implied by an unequal distribution of campaign contributions is not strong enough as compared to the degree of political inequality, which – in the context of the model – is needed to rationalize the voting outcome. One potential explanation could be that there are other sources of the wealth bias in the U.S., which are not captured by differential participation rates or campaign contributions.

---

[16]To construct this political Lorenz curve, I use campaign contributions shares along income groups from the Cooperative Congressional Election Survey from the years 2006 and 2008 (no data are available for 1980). These shares are computed by Bachmann and Bai (2013).

2.6   Quantitative exercise

2.6.1   Overview

In the quantitative exercise I simulate the dynamics of income inequality since 1980. Before doing so, I modify the environment of the benchmark model economy in two ways. First, I change the specification of the idiosyncratic employment process (section 2.6.2). This change is meant to account for a significant rise in wage dispersion in the U.S. during 1980-90. Second, I assume that agents anticipate the change in the employment process and vote once-and-for-all on the policy τ (section 2.6.3). This modification imitates the adoption of the Economic and Recovery Tax Act of 1981. Then I simulate the dynamics of income redistribution when the standard median voter and when the weighted median voter are decisive (section 2.7.1). To understand the differences between the voting outcomes under the standard and the weighted median voters, I quantify the effect of rising dispersion of wages and rising mobility across wage groups since 1980 on the most preferred tax rate of the standard and the weighted median voters (section 2.7.2).

2.6.2   Changes in employment process

Since 1981 the variation in wages in the U.S. started to increase dramatically. In figure 2.8 I plot the variance of log (male) hourly wages using the data from Heathcote et al. (2010a). Rising earnings inequality has been the main driving force behind growing inequality in income both at the bottom and the top of the market income distribution (figures 2.1-2.2). Different explanations have been proposed to explain this increased wage dispersion, such as globalization, decline of union strength, skill-biased technological change, executive compensation practices in the financial sector, etc. Since the rising wage dispersion itself is not the topic of this paper, I account for it by appropriately adjusting the parameters of the employment process.

The dispersion of productivities, measured by $\bar{\sigma}^2$ in the autoregressive representation, increases from 0.75 to 1.01. Second, the mobility across employment states rises: the persistence of shocks, measured by the autocorrelation coefficient ρ, decreases from 0.77 to 0.75. These numbers are the estimates obtained from the PSID data for 1995-96.

I assume that a regular household has the same probability of becoming super-rich, and a super-rich agent has the same probability of becoming a regular household as in the benchmark model economy. As a result, the long-run share of the super-rich in the total population doesn't change. Furthermore, I assume that the productivity level of the super-rich agent, ϵ_6, doesn't change.

The resulting transition matrix is shown in table 2.7 in Appendix. There is a larger dispersion in wages for regular households because ϵ_1, ϵ_2 (ϵ_4, ϵ_5) are lower (higher) than in

the initial benchmark model. Higher mobility comes from the fact that expected durations of the shocks for regular households are now lower than in the initial steady state.

2.6.3   Once-and-for-all voting in 1980

In 1980 agents vote once-and-for all on the rate of taxation τ, which becomes effective from 1981 onwards. Before voting, agents learn that the wage process has changed as described above. Since the identities of the standard and the weighted median voters are already known (section 2.5.3), what remains to be determined are their most preferred tax rates, $\hat{\tau}$. To find them, I repeat the steps of the computational strategy (Appendix) with slight modifications. In step 2, I feed into the model an updated process for the idiosyncratic risk. In step 4, I don't calibrate $\alpha^\star$ given $\tau_B^\star$, but instead find $\tau_B^\star$ given the already identified value of $\alpha^\star$.

The assumption of once-and-for-all voting is not as restrictive as it may appear first. In figure 2.9 I plot the annual estimates of the parameter τ.[17] The shaded region in the figure corresponds to the pre-voting period, in which the economy is assumed to be in the steady state politico-economic recursive equilibrium. It can be seen that in the period 1980-89 there was only one significant drop in the progressivity of the income tax rates (i.e. drop in the value of the parameter τ). This drop took place in 1981. Historically, this time corresponds to the adoption of one of the most important bills in the history of U.S. – the Economic and Recovery Tax Act. This Act drastically reduced the progressivity of the statutory income tax schedule. For example, the top marginal income tax rate fell down from 70 to 50 percent. The change in the statutory tax rates had a pronounced impact on income redistribution through taxes both at the bottom and the top of the income distribution (figure 2.3). No changes in the progressiveness of the tax code of the same magnitude occurred during 1980s.[18]

2.7   Main findings

2.7.1   Dynamics of income redistribution

---

[17]The estimates are taken from Gouveia and Strauss (1994). The authors estimated the tax schedule in (2.3) for the U.S. economy using individual level income tax data.

[18]The second, less significant, decline in the progressiveness took place in 1986, when The Tax Reform Act was adopted. The Act reduced the top marginal tax rates further down to 28 percent in 1988. At the same time, the same Tax Act eliminated the preferential tax treatment of capital gains, which explains a slight increase in progressiveness in 1987. The Tax Reform Act of 1986 seems to have had only short-terms effects in increasing the progressiveness, since the estimate for τ returns to its pre-reform levels shortly after 1986. Note that during the whole period the estimates for the parameters m_1 and m_2, which also affect progressiveness, didn't change significantly, see Gouveia and Strauss (1994). This allows me to evaluate the progressiveness by only looking at the estimates for τ.
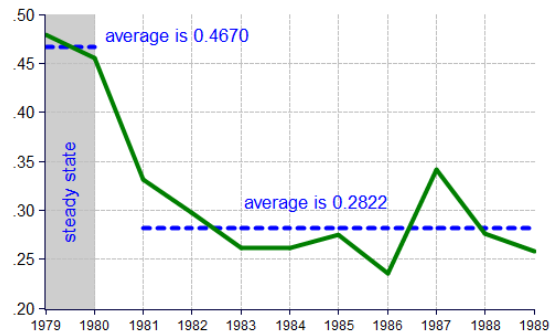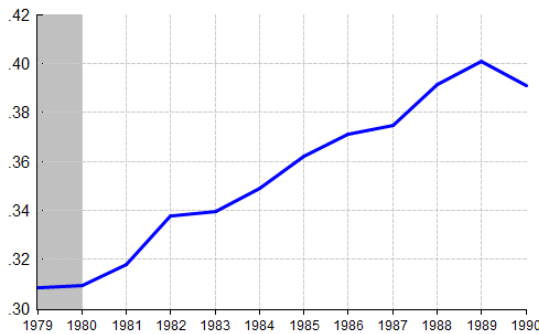
Figure 2.8: Variance of log (male) hourly wages

Figure 2.9: Estimates of the parameter τ

The equilibrium tax rates for the standard and the weighted median voters are shown in table 2.5. The column *Status Quo* refers to the benchmark model economy of 1978-79. The column *Total* shows the equilibrium voting outcomes under the changed wage process. The remaining two columns show the decomposition of this voting outcome due to mobility and inequality (see below). Thus, the model with a weighted median voter is able to account for roughly 36 percent of the decline in the tax rate from 46.7% to 28.2% during 1980s.

In figure 2.10, I plot the dynamics of market income inequality (first row) and disposable income inequality (second row) at the top (left column) and the bottom (right column) of the income distribution for the case of the standard and the weighted median voters.[19] On impact, both types of voters induce a jump in inequality in market incomes both at the top and the bottom. While inequality continues to rise under the weighted median voter, the standard median voter brings about a gradual decline in inequality below its original level in 1979. On impact, the disposable income inequality drops and continues to declines afterwards both at the top and the bottom under the standard median voter. The opposite is true for the case of the weighted median voter.

Figure 2.11 plots the overall redistributive effect of taxes and transfers (left column) and decomposes it into the redistributive effect of transfers (middle column) and the redistributive effect of taxes (right column). The first row shows redistribution at the bottom, while the second row shows redistribution at the bottom of the income distribution. The figure helps understand the dynamics of disposable income inequality in the previous graph. In particular, disposable income inequality mimics market income inequality under weighted median voter because the overall redistributive effect of taxes and transfers is almost zero. Under the standard median voter, the drop in disposable income inequality at the bottom is mostly driven by lump-sum transfers, whereas transfers and taxes are equally important in reducing inequality at the top. Overall, figure

---

[19]The dynamics are plotted until 1990 only because in the model the impact of voting on redistribution dies out around that time.

	Model				Data	
	Status Quo 1980	△ from mobility	△ from inequality	Total 1981-89	1980	1981-89
SMV	100%	−5%	+17%	110%	–	–
WMV	46.7%	+1.2%	−8.2%	40%	46.7%	28.2%

Table 2.5: The tax rate τ across experiments for the standard median voter (SMV) and the weighted median voter (WMV)
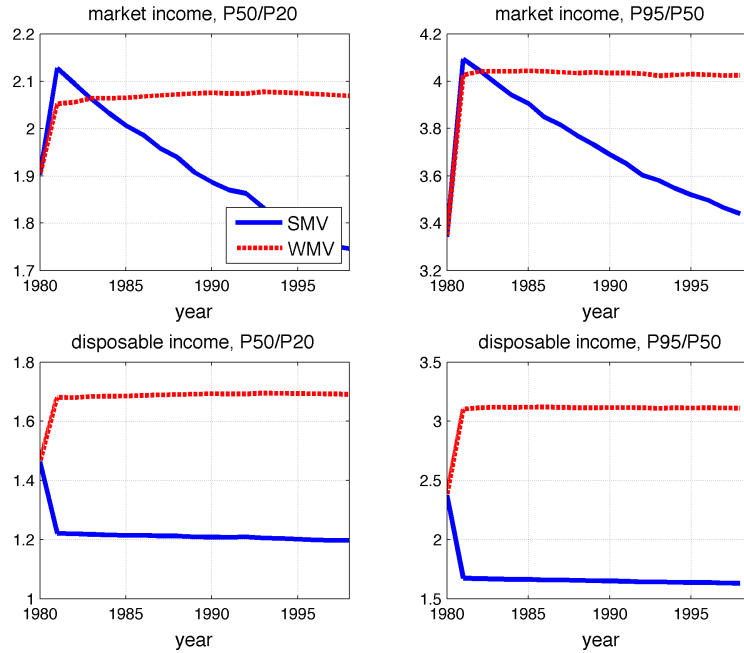
Figure 2.10: Dynamics of inequality in market income and disposable income at the top and the bottom of the income distribution in the model

2.11 suggests that a model with a weighted median voter predicts the dynamics of income redistribution quantitatively better than a model, in which the standard median voter is decisive.

2.7.2 Relative importance of mobility and inequality

In this section I quantify the effect of rising dispersion of wages as well as the rising mobility across wage groups on the most preferred tax rate of the standard and the weighted median voters in 1980. More specifically, I conduct the following counterfactual experiments:

1. In 1980 agents vote on the future rate of taxation anticipating that the dispersion of wages remains unchanged, while the mobility across wage groups increases (i.e. the levels of efficiency shocks correspond to the calibration in 1978 but the transition matrix corresponds to the calibration in 1995-96, table 2.7 on p. 57).
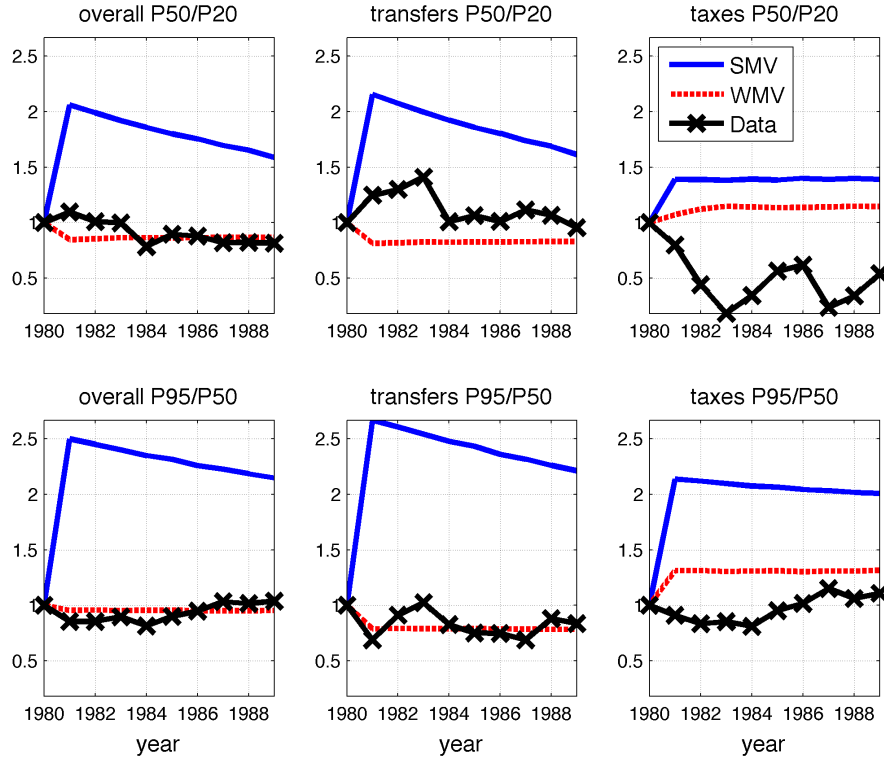
Figure 2.11: Dynamics of income redistribution in the model and in the data

2. In 1980 agents vote on the future rate of taxation anticipating that the mobility remains unchanged, while the dispersion increases (i.e. the transition matrix corresponds to the calibration in 1978-79 but the levels of efficiency shocks correspond to the calibration in 1995).[20]

The results of the experiments are presented in table 2.5. It can be seen that agents' expectations about future wage *inequality* account for a large portion of the change in the most preferred tax rates. Moreover, rising inequality has an opposite effect on the voting decision of both agents. The most preferred tax rate of the standard median voter increases by 17 percentage points relative to the status-quo. By the time of voting, this agent has accumulated a small amount of asset holdings and thus has built up a relatively low puffer against fluctuations in wages. The rising dispersion in wages increases the risk for this agent. The insurance effect pushes agent's most preferred tax rate up. The effect of rising dispersion is opposite for the weighted median voter, whose most preferred tax rate drops by 8.2 percentage points relative to the status-quo.

Increased mobility across wage groups also has an opposing effect on the most preferred policies of the two voters. The *prospect of upper mobility* shift the most preferred tax rate of the standard median voter downwards by 5 percentage points, while the prospect

---

[20]To conduct the experiments, I repeat steps 1-3 of the computational strategy described in Appendix. However, when computing the counterfactual tax rate in case of the bias (step 3), the strength of the political inequality in all the experiments is fixed at $\alpha^\star$, which I found in section 2.5.3.

of lower mobility make the weighted median voter reduce the tax rate by 2.2 percentage points.

2.8   Outlook

This paper is the first step towards understanding whether a wealth bias in the political process can improve on the result of the standard median voter in explaining the dynamics of redistribution in the U.S. But a lot more work needs to be done in the future.

First, we need a better understanding of the nature of the wealth bias in the political process. In this paper the bias is associated with agent's wealth according to an exogenously imposed rule. We need, however, a micro-theoretical explanation for how larger wealth translates into higher political power. Campante (2011) makes one step in this direction. In his framework the wealth bias arises endogenously, since the advantage of wealthier individuals in providing campaign contributions encourages parties to move their platforms closer to those individuals' preferred positions.

Second, as emphasized by Dahl (1961) in the epigraph of this paper, there might be other sources of the bias in the political process, such as knowledge (or education). In fact, in almost any empirical study on voter turnout the level of educational attainment is shown to have a significant and positive effect on the probability of casting a ballot.[21] Thus, people with higher education may exert stronger political power because they tend to go more often to the voting poll. Bourguignon and Verdier (2000) make a first step in studying this *educational bias* in politics. They construct a model, in which those at power (elite) are able to control the evolution of political institutions. This is done by *assuming* that political participation depends on the educational level of economic agents, and that the level of education is controlled by the elite. However, we still lack micro-theoretical foundations for the link between education and political participation.

Finally, it is important to understand how political inequality and income inequality interact *over time*. With once-and-for-all voting, changes in income and wealth inequality do not feed back to to the distribution of political power. Once sequential voting is introduced, changes in income and wealth inequality affect the distribution of voting rights. Then, rising wealth inequality implies a more skewed distribution of voting rights and therefore a more wealthy decisive agent. The decisive agent sets lower redistribution than her predecessor, which leads to an even higher wealth inequality. We can therefore study whether such an *inequality trap* is what drives political and wealth inequality in the U.S.

---

[21]Mueller (2003) gives an overview over the key empirical studies.

2.9   Appendix

Transition matrices

Productivity shocks ϵ					
0.27	0.52	1.0	1.91	3.66	300.07
Transition matrix Π					
0.5186	0.3704	0.1025	0.0081	0.0002	0.0003
0.1954	0.4290	0.3076	0.0639	0.0038	0.0003
0.0389	0.2394	0.4432	0.2394	0.0389	0.0003
0.0038	0.0639	0.3076	0.4290	0.1954	0.0003
0.0002	0.0081	0.1025	0.3704	0.5186	0.0003
0.05	0	0	0	0	0.95
Stationary distribution, %					
11.96	22.94	29.72	22.90	11.88	0.006
Expected duration, years					
2.07	1.75	1.79	1.75	2.07	20.00

Table 2.6: Transition matrix for 1978–79

Productivity shocks ϵ					
0.22	0.47	1.0	2.12	4.51	243.61
Transition matrix Π					
0.4999	0.3715	0.1167	0.0113	0.0003	0.0003
0.1975	0.4139	0.3101	0.0728	0.0054	0.0003
0.0445	0.2408	0.4291	0.2408	0.0445	0.0003
0.0054	0.0728	0.3101	0.4139	0.1975	0.0003
0.0003	0.0113	0.1167	0.3715	0.4999	0.0003
0.05	0	0	0	0	0.95
Stationary distribution, %					
12.0	22.89	29.74	22.85	11.92	0.006
Expected duration, years					
1.99	1.70	1.75	1.70	1.99	20.00

Table 2.7: Transition matrix for 1995–96

Computational strategy

1. I solve for the steady state recursive competitive equilibrium in 1979 using the employment process from table 2.6 and the model parameters from table 2.1. I obtain the stationary distribution of agents over the states $(a, \epsilon) \in \Gamma_1$ at $t = 0$, $f_0(a, \epsilon)$.[22]

2. For every $\tau \in \Gamma_2$, I compute the corresponding steady state recursive competitive equilibrium.[23]

---

[22]The set Γ_1 is a discretized version of the set $[0, \infty) \times E$.
[23]The set Γ_2 is a discretized version of the set $[0, 1]$.

3. I solve for transitional dynamics of the economy from the initial steady state computed in step 1 to each of the new steady states computed in step 2.[24] For each transition, I compute the corresponding function $v(a, \epsilon; F_0, \tau)$. I verify numerically that the function $v(a, \epsilon; F_0, \tau)$ is single peaked in τ for every $(a, \epsilon) \in \Gamma_1$. For every (a, ϵ), I find the most preferred tax rate $\hat{\tau}(a, \epsilon)$ according to (2.7).

4. Given the objects $\hat{\tau}(a, \epsilon)$ from the previous step and $f_0(a, \epsilon)$ from step 1, I find the equilibrium tax rate, $\tau_U^\star$, preferred by the standard median voter by setting $\alpha = 0$ in (2.6). Then I calibrate $\alpha^\star$, such that the weighted median of the ranked most preferred tax rates $\hat{\tau}(a, \epsilon)$ equals $\tau_B^\star = 0.467$.

---

[24]I use the guess and verify technique. It starts with an initial guess for the time path of the factor prices. Given this guess, the decision functions of the agent along the transition are computed. The initial distribution and the computed decision functions imply a new path for the factor prices. If the initial guess of the factor prices is different from this new path, the guess is updated accordingly. Domeij and Heathcote (2004) describe the procedure in detail.

Chapter 3

Voter mobilization and electoral competition

3.1 Introduction

The Republican congressional election campaign of 2002 showed the importance of voter mobilization. As it was typical of the previous federal election campaigns, the Republicans and the Democrats conducted massive advertising to persuade swing voters. But the Republicans also mobilized their potential supporters. Before the election, professional marketing researchers set up phone banks and door-to-door-canvassing operations to make a direct contact to the voters and identify their party affiliation. The county elections officials provided the voting history of registered voters, while third-party vendors provided data on demographic characteristics of the voters. All these data were then analyzed to identify those weak partisans who supported the Republicans but were likely to fail to go to the voting poll. The goal was to get them to the polls. Phone banks reminded people to register to vote and to vote on the election day, activists sign-waved at major intersections to show the way to the voting poll, rides were provided to the polling places.

Despite its potential effectiveness, voter mobilization is not widespread yet. As we document below, there is substantial heterogeneity in expenditures on voter mobilization, on the one hand, and advertizing, on the other hand, among candidates running in general elections for the House of Representative in 2010 and 2012 in the U.S. In this paper, we want to understand the driving forces behind this heterogeneity.

We suggest a simple model, in which candidates compete by running costly campaigns in order to increase their probabilities of winning the elections. Being agnostic about the choices of policy platforms, we focus on the campaign spending strategies of the candidates. There are two groups of voters in a district. The voters of the first type are partisan (core) supporters of the candidates. Such voters are already confirmed in

their decision for which candidate to vote in case they participate in voting. Voters of the second type are independent (swing). Such voters may vote for either candidate depending on electoral campaign impact and stochastic preference fluctuations. The voters of both types, however, may either participate in voting or abstain. Suppose that the participation decision is totally unrelated to the voters' political preferences and depends on some idiosyncratic random factors. During the electoral race candidates may affect swing voters' sympathies through advertizing and voter turnout through mobilization. We use this framework to illustrate several qualitative relations between equilibrium campaign spending and different characteristics of a given electoral district and given electoral race in general. A distinctive feature of our framework is that it allows us to analyze advertising and mobilization spending of candidates as parts of a single electoral campaign.

Our model reconciles persuasion and mobilization in a unified framework. Candidates act in two consecutive phases of the electoral race. In the first stage, they try to persuade independent citizens to vote for them by means of costly advertizing activities. In the second phase, candidates mobilize their core supporters after receiving a coarse signal about the state of relative popularity which has been realized after the advertizing stage. Since both advertizing and mobilization activities are costly, candidates also exert fundraising efforts which are the source of disutility in their objectives.

In accordance with the intuition from probabilistic voting models, we assume that there is a stochastic factor (popularity shock), which may affect voters' preferences and, as a result, the voting outcome. We suggest a novel way of modeling such a stochastic shock. Namely, we use Markov chain transitions between the discrete states of candidates' relative popularity. Such discrete states are associated with the coarse information state-space of candidates about the fractions of their supporters among independent citizens. The assumption of the coarse discrete state-space (popularity grid) stems from natural limitations in acquiring precise information about the exact split of voters' latent sympathies. Thus, instead of getting a precise value of the sympathy split, candidates are aware only about the closed interval from the discrete partition, where the latent split is currently located. In turn, a Markov transition matrix reflects the candidates' subjective probability assessments for stochastic switching between these discrete popularity states. During the campaign, candidates receive coarse information about the current state of popularity and coordinate their spending strategies conditional on this information.

From the analysis of our novel modeling framework we derive a set of qualitative results. The first qualitative finding is that the amounts of advertizing and mobilization expenditures depend differently on the expected turnout rate and on the fractions of swing/partizan voters in a district. In particular, at the intensive margin candidates advertize more and mobilize less as long as the expected participation rate and the fraction of swing voters are higher. These findings are quite intuitive and reflect the fact that these characteristics of the electoral environment affect the relative efficiency and, hence, the

marginal effects of advertizing and mobilization activities. So, advertizing has a larger marginal effect if the fraction of independent voters is higher. At the same time, the fraction of partizans is lower, which means that their mobilization has a lower effect on the voting outcome. In contrast, if expected turnout is low, mobilization is getting more important due to the larger fraction of core supporters who can be mobilized for voting, but advertizing is getting less effective since most of persuaded swing voters will simply not vote. Note, that in case candidates have fixed campaign budgets, such reasoning would determine an optimal split of campaign funds for advertizing and mobilization expenditures.

The second qualitative finding establishes the link between equilibrium spending and the degree of ex-ante electoral advantage of one of the candidates, which can be interpreted as, for example, an incumbency advantage. Both advertizing and mobilization expenditures decrease as far as an ex-ante disparity between candidates rises (expected competition is low). Such a qualitative result is not new and is in line with the findings by Snyder (1989), Erikson and Palfrey (2000), Herrera et al. (2014), etc. Nonetheless, it is generated by the framework, where both advertizing and mobilization are parts of the same campaign and this delivers some new intuition. In the cited papers, such a competition effect (positive relation between spending amounts and the degree of competitiveness) usually comes from the properties of the contest success function (see Rosen (1986)), which exogenously maps candidates' efforts to the electoral outcome. In our case we don't refer to specific functional assumptions and capture the competition effect by the more intuitive reasoning. As for advertizing, an ex-ante advantage of some candidate implies the presence of external factors (like administrative resource, higher valence, more appealing platform, etc.), which contribute to the split of voter sympathies towards the leader candidate. Thus, the relative role of persuasion is getting smaller (for both candidates) as compared to the case in which there is no clear leader/underdog and the external factors are balanced. As for mobilization, the competition effect is revealed in expectations. Naturally, if one candidate is an ex-ante leader, then it is more likely that she remains a leader after the advertizing stage, as well, and will have relatively more supporters among swing citizens. In that case, candidates spend less on mobilization, since it is less likely that mobilized partizans will be pivotal in determining the winner on the voting day as compared to the state, where candidates have equal constituencies.

We test some of the implications of our model. We use the administrative data on campaign expenditures by the Federal Election Commission for the 2010 and 2012 House of Representative election cycles. Using the Tobit regression model, we confirm that both advertizing and mobilization expenditures of candidates increase with expected competitiveness of the race, where we control for the expected competitiveness using the *2010 and 2012 House Race Ratings* by *The Cook Political Report.* We also confirm that the expected turnout, which we proxy by voter turnout rates in the previous election in

the same congressional district, has a positive effect on advertizing expenditures, while we don't find any significant effect of past turnout rates on mobilization expenditures.

We contribute to the literature on electoral campaign spending (Snyder (1989), Nalebuff and Shachar (1999), Herrera et al. (2014)). We consider implications of different characteristics of the electoral environment on the equilibrium spending strategies of candidates. In contrast to the aforementioned papers, advertizing and mobilization are analyzed in a united framework as two consecutive phases of the same electoral race. Our general findings regarding advertizing and mobilization campaign spending are mainly in line with the corresponding findings from the literature, where these activities have been considered separately.

As for methodology, our modeling framework contributes to the class of probabilistic voting models (Lindbeck and Weibull (1987), Lindbeck and Weibull (1993)). We introduce an original and intuitively appealing mechanism of a stochastic change in voter preferences which is a Markov chain transition between the states of candidates' relative popularity. Such mechanism naturally implies a discreteness of the switching variable. It, therefore, is in line with a realistic assumption that candidates have only limited knowledge about the current state of their relative popularity. They may infer only a coarse information from a discrete grid of relative popularity states, but not the exact value of the fraction of their supporters. Also, it is naturally to assume that candidates implicitly use a discrete transition model to make a probability assessments about their relative popularity in the future.

Our empirical strategy is close to Abramowitz (1991) and Erikson and Palfrey (2000), who study the effect of campaign spending on candidate's vote share. Both papers tackle the simultaneity problem, which arises, when one regresses the vote share on campaign spending. Erikson and Palfrey (2000) construct a game theoretic model to show formally that the simultaneity bias is minimal in close elections. They confirm their finding using the House races in 1974-80 and 1984-90. Abramowitz (1991) uses as a measure of electoral expectations the *Congressional Quarterly*'s assessments of the expected vote. Both of these papers are agnostic about the nature of campaign spending, while we show – theoretically and empirically – the importance of mobilization and advertizing expenditures. We argue that the *Congressional Quarterly*'s assessments are very narrowly defined and not truly endogenous to campaign spending. Instead, we use a better measure of expected electoral closeness.

We proceed as follows. In section 3.2, we document two empirical facts: large heterogeneity in voter turnout rates and campaign spending on voter mobilization across congressional districts. In section 3.3, we describe the modeling environment for electoral competition where two candidates run their campaigns and aim to win the office. We start with a brief model sketch, give general intuition and discuss the simple mechanisms

which affect candidates' decisions regarding their advertizing and mobilization spending. We then briefly analyze equilibrium spending strategies through the lens of the suggested framework. These equilibrium spending patterns make a ground for the testable hypotheses that we test in section 3.4. Section 3.5 draws conclusions.

3.2 Empirical facts

Empirical fact 1: Heterogeneity in voter turnout rates

We first document that there is substantial heterogeneity in voter turnout rates across congressional districts during the House of Representative elections in the U.S. To illustrate this fact, we first compute the turnout rate in each congressional district during the 2012 House elections. Then we average turnout rates across all congressional districts within each state. Figure 3.1displays average turnout rates by state. Appendix 3.6.1 contains all the details regarding the data. If we look at the 2010 House elections, there is a qualitatively similar pattern of heterogeneity in turnout rates, though the overall level of the turnout is even lower than in the 2012 election cycle.
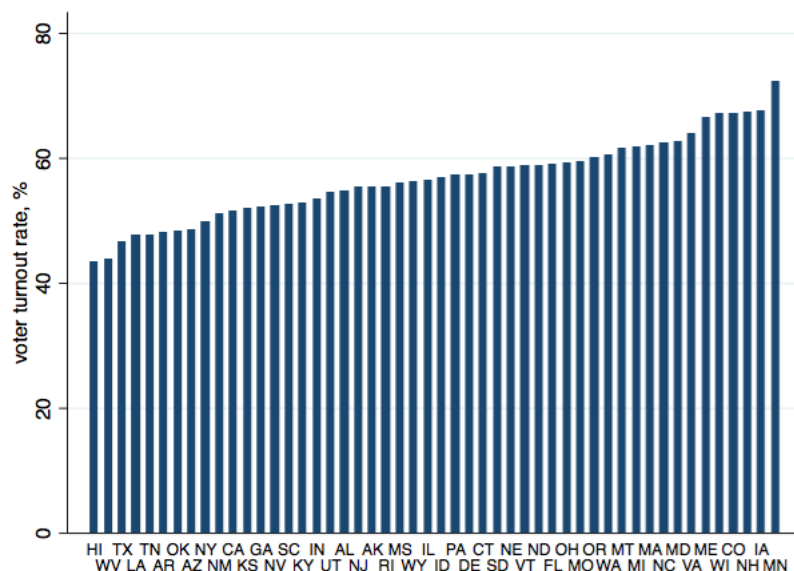
Figure 3.1: Voter turnout rates during the House elections of 2012 averaged across congressional districts, by state

As can be seen from the figure, voter participation ranges from as low as 43 percent in Hawaii (HI) to no more than 72 percent in Minnesota (MN), while the average turnout rate is 55 percent. The nation's most populous states – California (CA), New York (NY) and Texas (TX) – rank in the very bottom.

Several factors may help explain the heterogeneity in turnout rates. First, traditionally active citizenry explains high turnout rates in some states such as Minnesota.

Minnesota has been ranked first in turnout rates in eight of the last nine mid-term and presidential elections. Second, simplified legislation on voter registration explains high turnout rates in Iowa (IA), Maine (ME), Minnesota, New Hampshire (NH) and Wisconsin (WI). This legislation allows voters to register or update their registration at the polls on the election day. In contrast, the states with lowest turnout rates – Tennessee (TN), Arkansas (AR), Texas, Oklahoma (OK), and West Virginia (WV) – have more burdensome registration requirements. North Dakota (ND), ranking among the states with highest turnout, is the only state, which doesn't require registration. Third, the composition of the voting age population, which is the denominator in our definition of the turnout rate, matters. Growing share of Hispanic electorate, who vote at lower rates than other population groups, explains lower turnout rates in the southern states, i.e. Arizona (AZ), Oklahoma, Texas, and Louisiana (LA). Finally, voter mobilization campaigns boost participation of voters. As we will show below, some candidates in Texas, Louisiana, Nevada (NV), Delaware (DE) and New Jersey (NJ) massively directed their campaign funds into getting out the votes of their potential supporters.

Empirical fact 2: Disbursements on voter mobilization

Previous literature has emphasized the importance of candidate's expenditures on voter persuasion or advertising. In this section, we focus on candidate's expenditures on voter mobilization. We use the data by the U.S. Federal Election Commission on disbursements on voter mobilization and advertising, reported by candidates running in the general election for the House of Representatives during the election cycles of 2010 and 2012. For each candidate, we compute the share of expenditures on mobilization in the total expenditures, where we define the total expenditures as the sum of disbursements on advertising and voter mobilization. All the details regarding the data can be found in appendix 3.6.2.

Several findings emerge from studying the data. First, a relatively large fraction of candidates prefers not to invest in mobilization activities altogether. In our sample comprising of 1,463 observations, more than 27 percent of the candidates virtually invest in advertising only (i.e. have zero shares spent on mobilization). Second, at the intensive margin there is substantial variation in the expenditures on mobilization. Figure 3.2 shows the distribution of candidates' expenditures on mobilization averaged across congressional districts by state for those candidates, who spend strictly positive amounts of campaign funds on mobilization. In the figure we pool together the 2010 and 2012 election cycles but the results are quantitatively similar if we consider each election cycle separately.

In 13 states the average share of mobilization disbursements is below 10 percent: e.g. Minnesota, Iowa, Oklahoma and Connecticut (CT). In roughly half of the states mobilization expenditures range between 10 and 20 percent (e.g. Arizona, California,

Wisconsin), while in 12 states candidates spend on average more than 20 percent of their campaign funds on mobilization (e.g. Wisconsin, Nevada, Delaware and New Jersey). Vermont (VT) is a clear outlier with an average share of more than 45 percent spent on mobilization.
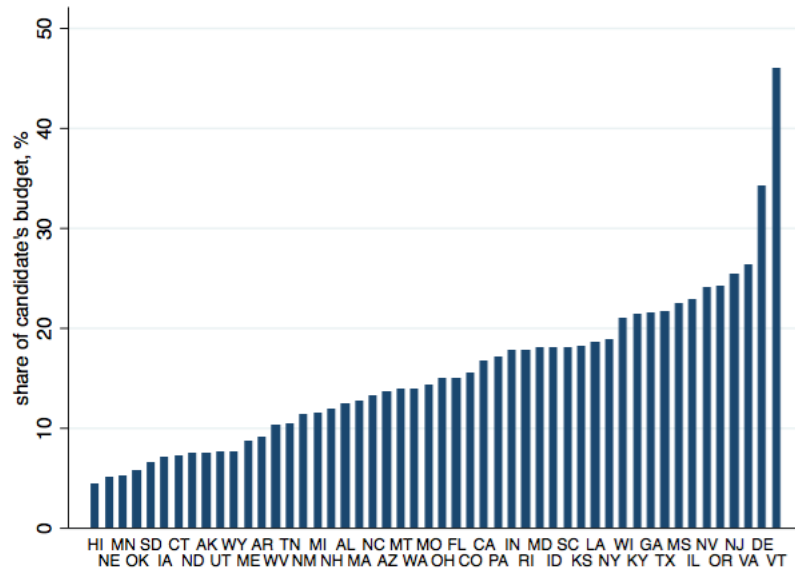
Figure 3.2: Share of candidate's budget spent on mobilization during the House elections of 2010 and 2012 averaged across candidates and congressional districts, by state

3.3   Model

3.3.1   Sketch of the Framework and Intuition

Consider a situation, in which two candidates compete for a single chair in a legislative body in a given electoral district. Candidates run their campaigns and aim to win a chair in winner-takes-all elections. If one candidate gets more than half of total ballots casted in the actual voting, such a candidate wins.

Suppose that all eligible voters in a district may be of two types. The voters of the first type are partisan (core) supporters of the candidates. Such voters are already confirmed in their decision for which candidate to vote in case they participate in voting. Voters of the other type are independent (swing). It means that such voters may vote for either candidate depending on the electoral campaign impact and the stochastic preference fluctuations. The voters of both types, however, may either participate in voting or abstain. Suppose that the participation decision is totally unrelated to the voters' political preferences and depends on some idiosyncratic random factors.

Candidates naturally pursue a dual goal in the course of the electoral campaign. On the one hand, they try to affect voters' sympathies by means of advertizing (persuasion).

Since only swing voters may change their political sympathies during the electoral race, persuasion is intended for them. On the other hand, candidates aim to increase the actual voting participation of their supporters through voter mobilization activities. It is natural to assume that partisan voters are the main target of mobilization because of several obvious reasons. Firstly, only partisans can be considered as firm supporters of a given candidate. Thus, mobilizing partisans, in contrast to mobilizing independent voters, is safe and can't lead to erroneous mobilization of those swing voters who may vote for an opponent. Secondly, political preferences of the core supporters are supposed to be persistent and reflected by voter surveys and pools carried out long before the ongoing race. As a consequence, it is easier for a candidate to identify her partisans rather than to search for supporters among swing voters whose sympathies are much more floating, harder to detect and may alter during the ongoing race.

For expositional simplicity and to focus on candidate's behavior, we treat voters as static responders, rather than rational agents. Thus, both advertizing and mobilization technologies are modeled as exogenous mechanisms, which map campaign spending into popularity or turnout. However, these mechanisms can be rationalized along the lines of existing models, where voters make their decisions by rationally comparing costs and benefits of their actions.

By conducting advertizing and mobilization activities, candidates attempt to increase the number of ballots casted in their favor in the actual voting. Important to note that in reality electoral outcome is not determined solely by candidates' campaign efforts but is also affected by many unanticipated incidents and random factors. Thus, the voting outcome is generally probabilistic and candidates' campaign activities are aimed to affect the *probabilities* of winning the elections. Following these lines, we assume that the effects of both advertizing and mobilization activities are not purely deterministic conditional on candidates' spending.

Furthermore, as it is generally assumed in the probabilistic voting models, we suppose that there exists an aggregate shock that influences sympathies of swing voters. Such a shock may be interpreted as a political scandal or a sudden change in the state of the world, which significantly changes voter preferences. The technical purpose of this shock in the modeling framework is to guarantee a non-deterministic voting outcome for any starting positions of candidates and model parameters. We introduce such a popularity shock in a manner that is not standard in the related literature, but we argue that this way of modeling is quite realistic, intuitive and convenient for the analysis and equilibrium characterizations.

More particularly, we assume that candidates are able to discover their relative popularities among independent voters only with a limited precision up to a set of discrete states. Such an assumption is pretty natural and doesn't contradict the reality. As an

example, we can imagine that candidates hold a sociological survey (that has a limited precision) during the electoral race and receive the outcome according to which candidate A has $45 - 48\%$ of support and candidate B has $52 - 55\%$ of support respectively. Conditional on this information, candidate A knows that her fraction of supporters among non-partizan citizens is distributed[1] between 0.45 and 0.48. Thus, candidates associate their instant popularities with a discrete interval (state) from some partition of $[0, 1]$, rather than with a precise value from $[0, 1]$. We assume further that the stochastic shock of popularity switches (probably, more than once) these discrete states of the relative support according to a Markov chain process in the course of electoral race. Thus, candidate A understands that on the voting day her relative popularity may change from $[0.45, 0.48]$ to some other state from the discrete partition, for example $[0.48, 0.52]$, with some (commonly known) transition probability.

Stochastic transitions across the discrete states of relative support directly interfere with the campaign spending strategies of candidates. While doing advertizing candidates attract the sympathies of swing voters and, thus, affect the location on a discrete popularity state-space. Hence they also affect the probability distribution of terminal (voting day) popularities through the Markov chain transition matrix. When candidates mobilize, they increase the expected number of actual votes of their partizans. However, depending on the expected terminal state of popularity, the voter mobilization may be more or less likely pivotal for the election outcome. As a result, our popularity shock suggests a stochastic mapping from the candidate campaign efforts on the probabilities of winning the elections. We emphasize that such a mapping takes a simple and convenient analytical form and we illustrate it further in this section.

3.3.2 Basic Model

We now formalize the ideas written above. We describe a simple model of electoral competition, where two candidates compete in an electoral district for an office. In our framework we omit the part of electoral competition, in which candidates choose their political platforms, by assuming that politicians have already committed to some programs. Instead, we focus on the campaign spending side of the race. Formally, the game consists of two periods. Candidates run their electoral campaigns by firstly spending on advertizing and then on mobilization. Therefore, in order to increase the winning chances they affect not only voters' preferences but also voters' turnout.

Electoral district. In what follows, index $j \in \{A, B\}$ will refer to the one of two candidates. Electoral process is organized in a district with a unit mass of citizens

---

[1]It may be assumed that the distribution is uniform. Although this may seem as a simplification, it is reasonable assumption especially in case of electoral units with not too large electorate constituencies, where the Law of Large Numbers and the Central Limit Theorem work just approximately well.

eligible to vote. We suppose that citizens in our model are a part of environment rather than strategical players (as in Snyder (1989), Erikson and Palfrey (2000), for example). They respond to the electoral campaigns of the candidates and to the stochastic shocks. The fraction $1 - \mu$ are swing voters who can be persuaded by means of the advertizing campaign. Each candidate has a fraction $\mu/2$ of core supporters (partizans) who can be mobilized to vote. The candidate who gets more than half of total ballots wins and receives the winning rent $R > 0$.

Fundraising. Campaign activities are costly for candidates. We may think of associated costs as fundraising efforts or direct efforts of campaigning. These efforts contribute disutility into the objectives of candidates. For simplicity, we suppose that such disutility is a linear function of spending amounts, so each unit of campaign funds costs $c > 0$.

Advertizing. The campaign spending game starts with the advertizing phase. In our model advertizing is a direct competition between candidates for the sympathies of swing voters. By doing advertizing politicians aim to increase their expected constituencies among independent voters. The strategy of candidates is to choose simultaneously and independently positive amounts x_j that they firstly raise and then spend in the course of advertizing part of campaign. In literature the advertizing competition is usually modeled through some mapping from the campaign efforts of candidates onto the set of possible preference splits of electorate, or directly onto the set of winning probabilities. For instance, Snyder (1989), Skaperdas and Grofman (1995) used a contest success function (Tullock (1980), Rosen (1986)), whereas Grossman and Helpman (1996) used an implicit functional form, where the difference between spending of two candidates is the argument. We will follow the latter option and suggest an implicit mapping from candidates spending to the fraction of voters supporting the given candidate. We also augment our advertizing competition with the random factor.

The candidate A receives the fraction of swing voters given by:

$$
s^0 = \begin{cases} 1 & \text{if } \frac{1}{2} + H(x_a, x_b; \Delta) + u > 1 \\ \frac{1}{2} + H(x_a, x_b; \Delta) + u & \text{if } \frac{1}{2} + H(x_a, x_b; \Delta) + u \in (0, 1) \\ 0 & \text{if } \frac{1}{2} + H(x_a, x_b; \Delta) + u < 0 \end{cases} \tag{3.1}
$$

where H is $\mathbb{C}^2$ function from $\mathbb{R}_+^2$ to $[-\frac{1}{2}, \frac{1}{2}]$, strictly increasing and concave with respect to the first argument and such that $H(x_a, x_b) = -H(x_b, x_a)$. Parameter Δ corresponds to the mass of persuaded swing voters and reflects the scale effect of advertizing spending. Namely, we require that $H'_x(\Delta)$ is a decreasing function of Δ. It implies that advertizing is getting less efficient as more voters are advertized. This intuition is in line with the theories of informative campaigning (see Schultz (2007), for example).

Variable $u \sim U[-\frac{1}{2\varphi}, \frac{1}{2\varphi}]$ is a random component which introduces uncertainty into

the outcome of the advertizing stage. The intuition behind this random component is quite straightforward. Although the advertizing spending increases the chances of getting a larger fraction of supporters, there are a lot of other issues related to persuasive campaigning, which may affect the sympathies of the electorate. For example, the choice of advertizing agenda that may or may not match the interests of voters, performance on the debates, noisy effects of negative advertizing as well as many other aspects make the outcome of advertizing campaign far from being perfectly predictable. As a result, candidate A gets the fraction s^0 of sympathized swing voters and candidate B receives $1 - s^0$.[2]

Information state-space of candidates. After the advertizing phase ends, candidates receive a coarse signal about the realized fraction s^0. We suppose that they can't infer the precise value s^0 due to natural limitations of the pooling and survey precision. Instead, they get an information that s^0 belongs to a particular range (state) S^0. The discrete information state-space of candidates is limited up to the following partition of $[0, 1]$:

$$S^0 \in \{S_b, S_c, S_a\} \equiv \left\{ [0, \frac{1}{2} - \alpha), [\frac{1}{2} - \alpha, \frac{1}{2} + \alpha], (\frac{1}{2} + \alpha, 1] \right\} \qquad (3.2)$$

where α is some exogenous value from $(0, \frac{1}{2})$. Intuitively, state S_a tells that candidate A is ahead in the race, state S_b suggests that candidate A is behind and state S_c informs that the race is close at the instant moment and both candidates have approximately equal support. Certainly, the number of states in the partition may be significantly higher. We here focus on the simplest case with just 3 states to deliver the main implications of the framework.

The variance of advertizing error is large enough, namely $\varphi \leq \frac{1}{1+2\alpha}$.

If parameter φ is sufficiently small, then the information that state $S^0 = S_c$ has been realized means for candidates that s is uniformly distributed in the interval defined by S_c. It is easy to derive that the advertizing mechanism introduced above implies the following probability distribution of states S^0 conditional on candidates' spending

$$\bar{\pi}_{S^0} = \begin{pmatrix} \frac{1}{2} - \varphi\alpha - \varphi H(x_a, x_b; \Delta) \\ 2\varphi\alpha \\ \frac{1}{2} - \varphi\alpha + \varphi H(x_a, x_b; \Delta) \end{pmatrix} \qquad (3.3)$$

---

[2]We note that the mechanism of preference formations implicitly incorporated into (3.1) can be readily rationalized in the spirit of spatial probabilistic voting models (Lindbeck and Weibull (1987), Lindbeck and Weibull (1993)). In particular, we may suppose that idiosyncratic preferences of the swing voters are distributed on some one-dimensional preference space and candidates bring some aggregate preference shift by campaigning. However, since the intuition doesn't change at all in that case, we keep our simple implicit function and let it determine the outcome of advertizing competition in response to candidates' campaign spending.

where the top entry contains the probability of having state S_b, the middle entry - S_c and the bottom entry - S_a.

Advantage. So far all assumptions in our framework suggested symmetry between two candidates. We now introduce asymmetry, which can be interpreted as an ex-ante advantage of one candidate over the other. In doing so, we follow the principles of the campaign spending and valence accumulation games in the literature.[3] In general, the logic is following. Suppose that candidates spend equal amounts of money (exert equal levels of effort) in the campaign. Then, if a candidate has some ex-ante advantage, her expected number of votes is higher than that of the opponent. In the literature, it is usually interpreted as incumbency advantage, but may be also considered as advantage in relative popularity of political platforms or in valence.

In our basic model we assume that this advantage is reflected in the tastes of independent (swing) voters.[4] Without loss of generality, assume that candidate B is a leader (i.e. she has some advantage). We assume that after candidates observe S^0, an additional stochastic effect related to the advantage of candidate B over A takes place. In particular, with probability $\gamma \in [0, 1]$ state S^0 switches to state S_b and with probability $1 - \gamma$ it doesn't change. We denote the state realized as a result of this advantage effect by S'. Candidates observe S' after its realization. Since candidates make no actions between receiving signals S^0 and S', we may equivalently think that the advertizing outcome and the advantage effect realize simultaneously, but independently of each other. The probability distribution of states S' conditional on advertizing spending, compared to (3.3), is getting additionally shifted towards the state S_b, where candidate B is more popular,

$$\bar{\pi}_{S'} = (1 - \gamma)\bar{\pi}_{S^0} + \gamma \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{3.4}$$

Parameter γ determines the degree of an advantage of candidate B. If $\gamma = 0$, there is no difference between candidates. As γ is getting higher, the relative advantage of candidate B is increasing. In case $\gamma = 1$, an advertizing competition has no meaning and $S' = S_b$ with probability one. In terms of our modeling framework, the advantage of a candidate implies that, *ceteris paribus*, she is more likely to receive higher support among

---

[3]See Snyder (1989), Grossman and Helpman (1996), Meirowitz (2006), etc.

[4]Of course, in reality the advantage may be captured by many other factors. For example, candidates may differ in costs of fundraising, in the efficiency of advertizing and mobilization technologies, in shares of core supporters, etc. A careful study of the effects of all these distinct advantages on the electoral competition requires a separate analysis and we will not provide it in the current paper. We note, however, that some of these alternative advantage types lead to similar qualitative implications.

independent voters than her opponent.[5]

Stochastic shock of popularity. We introduce another randomness into the framework that conceptually is closely related to the aggregate shock of popularity in probabilistic voting models. Intuitively, it captures the idea that some factor, which is orthogonal to advertizing campaign of candidates, may perturb the preferences of voters. For example, some sort of scandal may happen with one of the candidates and the sympathies of electorate may switch dramatically afterwards. Also, we may think that the voter preferences are state contingent and if the "state of the world" changes (e.g., a terrorist attack happens), voters may revise their political sympathies. From the technical point of view, such a shock allows us to avoid ex-ante deterministic voting outcomes in the winner-takes-all type of elections. In our framework we assume that this popularity shock affects the preferences of swing voters only, but not those of the partisans.

We suggest a non-standard specification for the stochastic shock to accommodate the assumed discrete information field of the candidates. Namely, we suppose that the shock of popularity is a Markov chain switching from a state $S' \in \{S_b, S_c, S_a\}$ to a state $S'' \in \{S_b, S_c, S_a\}$, where S'' denotes the popularity state realized on the voting day. The conditional probabilities of getting into some state S'' from some state S' are determined by a transition matrix P (which, in our case, is a 3×3 stochastic matrix). A generic element of P can be defined as

$$P_{S',S''} = prob\left(s'' \sim U(S'') | s' \in S' \right) \tag{3.5}$$

where s' and s'' are latent fractions of swing voters supporting candidate A formed prior to the popularity shock and in result of the popularity shock, respectively. Notation $s \sim U(S)$ means that s is distributed uniformly on the interval associated with state S. We assume that transition probabilities from P are common knowledge.

There are several reasons for why the Markov chain assumption is attractive. First of all, it naturally fits the idea of discrete perception of voter sympathies by candidates and suggests a convenient probability assessment over such discrete state space. Secondly, this specification of the stochastic shock is pretty convenient for our analysis, since many quantities of interest may be easily expressed using the matrix notation. Finally, specified in a proper way, the transition matrix P is able to generate a realistic persistence in transitions from one state to another. Intuitively, it means that if initial state S' suggests that candidate B is a leader, than it is more likely that in the following state S'' candidate

---

[5]It is important to note that the advantage naturally plays the role of a stochastic shift of a latent support fraction s^0 towards 0 (so the leader B gets more support). Alternatively, an advantage parameter could be introduced directly into the advertizing function (3.1) and give a similar effect. Nonetheless, we prefer our separate formulation of an advantage effect. Therefore, we don't relate an advantage directly to the advertizing technologies and provide a broader field for possible interpretations. In addition, introducing an advantage parameter directly into the probability distribution of discrete states is technically more convenient than map it to this distribution from (3.1).

B will remain to be a leader. This feature is not captured in most of probabilistic voting models, since the popularity shock there is usually assumed to be continuous and uniform, so all feasible realizations of the preference split after the shock are equally likely.

For expositional simplicity we assume the following parametrization of P:

Transition matrix P depends on just one parameter, ρ, as follows

$$P = \begin{pmatrix} 1 - 2\rho & \rho & \rho \\ \rho & 1 - 2\rho & \rho \\ \rho & \rho & 1 - 2\rho \end{pmatrix} \tag{3.6}$$

and $0 < \rho < \frac{1}{3}$.

This assumption implies that the shock randomly shifts the popularity states symmetrically with respect to the initial state S'. Also, such a transition exhibits "memory", since the probability of staying in the same state is higher than the probability of shifting to any other state.[6] The latter property is significant, since it emphasizes the importance of the advertizing stage because advertising directly affects the realization of S'.

Abstention from voting. The assumption of a continuum mass of citizens in the district naturally implies absence of pivotality. Hence, citizens are indifferent about voting in elections if nothing else is assumed. We assume the following (for all types of voters): (1) the benefit of voting $b \geq 0$ is the same across all citizens and doesn't depend on the favorite candidate, candidates' characteristics, or its popularity; (2) the cost of voting c_i is an exogenous non-negative random variable that is identically distributed, independently realizes across citizens and determines the cost associated with participation in elections; (3) a citizen i votes if $b \geq c_i$ and abstains from voting otherwise.

Together these assumptions imply that the participation decision is orthogonal to the political preferences and sympathies of voters towards the candidates. Therefore, the advertizing stage in our model doesn't affect the voter turnout.[7] The benefit of voting b can be associated with social responsibility, civic duty or some other ethical considerations. For simplicity, we take b to be the same among all citizens, although it is not necessarily the case in reality. Since the participation decision depends on the *difference* between benefits and costs, it is sufficient to only assume heterogeneity in costs to achieve heterogeneity in participation. Suppose the support of the cost distribution is $[0; +\infty)$ and denote the corresponding CDF by Φ. Thus, Φ determines the natural turnout rate $\delta = \Phi(b)$. By the natural turnout we mean the share of citizens, who vote

---

[6]We note, that a sort of asymmetry between candidates can be introduced into the model through the parametrization of P. In particular, we may suppose that transition probabilities are tilted towards state $S'' = S_a$ or $S'' = S_b$, indicating an expected advantage in popularity of candidate A or B respectively. Alternatively, an advantage effect can be represented as an additional biased popularity shock with its own transition matrix.

[7]This assumption agrees with recent empirical findings in Spenkuch and Toniatti (2015).

in the absence of any mobilization efforts of the candidates.

Voter Mobilization. After observing the realization of the state S', candidates step into the mobilization phase of their campaigns. We model mobilization as a tool that candidates use to increase the voter turnout among their supporters. By assumption, candidates may detect only those supporters who belong to the group of partisans of a particular candidate and can't strongly identify supporters from the pool of swing voters. Hence, candidates induce *group mobilization* increasing the turnout only among the core supporters, thus implementing mobilization targeting.[8] The candidates at this stage choose simultaneously and independently positive amounts y_j to raise and spend for mobilization of core supporters. The key issue regarding this choice is that the marginal effect of mobilization on the probability of winning depends on the unknown state of popularity S'' on the voting day. Thus, mobilization spending for a candidate is similar to buying an asset with state-contingent payoffs.

Similar to the case of advertizing, we accept a simple implicit mechanism which maps mobilization spending to the turnout of partizans. At the same time, this mechanism implies a stochastic outcome of the mobilization effort. In particular, we suppose that the turnout of partizans is a Bernoulli random variable with the success probability depending on the mobilization effort of a candidate. Thus, with probability $q(y_j)$ all partizans of candidate j are mobilized and with probability $1 - q(y_j)$ the mobilization effort has no effect on the turnout of partizans. We assume that function $q(y)$ is $\mathbb{C}^2$, strictly increasing and concave, $q(0) = 0$ and $\lim_{y \to \infty} q(y) = 1$. Therefore, if mobilization spending is higher, it is more likely that a candidate induces full participation of her core supporters.

Such a minimalistic mobilization mechanism is chosen to simplify the analysis without shifting the focus from the qualitative implications of the strategic mobilization campaigning. Note, that conditional on spending amount y_j, the expected fraction of j's partizans who participate in the election is

$$E\delta_j(y_j) = \delta + (1 - \delta)q(y_j), \qquad j \in \{A, B\} \tag{3.7}$$

Thus, we may interpret $q(y)$ as a factor that determines the *expected* turnout rate of partizans. It then allows to rationalize our mobilization function and agree it with the logic of the mobilization models in the literature.[9] Namely, we may suppose that candidates

---

[8]Of course, in reality candidates not necessarily mobilize only their partizans. By inducing canvassing and pooling activities, candidates may identify supporters among independent voters and mobilize them as well. Although the general intuition is not completely different in that case, it is an interesting extension which implies a closer interaction between advertizing activities ("fighting" for supporters) and mobilization phase (getting them out to vote). We refer more close inspection of this possibility to future research.

[9]See, for example, Nalebuff and Shachar (1999), Herrera et al. (2008), Herrera et al. (2014), etc.

use mobilization spending y to increase the voting benefit b of partizans (buying votes) by uniformly spreading y among them. Then, an expected turnout rate will become $E\delta(y) = \Phi(b + \lambda y)$, where λ is a coefficient (possibly depending on other model parameters), which reflects scale effects (an amount of mobilized people), effectiveness of spending, etc. By relating it to (3.7), it may become possible to approximate function $q(y)$ given a properly chosen Φ.[10]

From the technical point of view, the Bernoulli assumption allows us to determine the range of latent values s'', where mobilization can be pivotal with non-zero probability, i.e. can change the voting outcome conditional on the terminal split of swing voters between the candidates. It can be shown that such a range is given by $s'' \in (\frac{1}{2} - \beta, \frac{1}{2} + \beta)$, where

$$\beta = \frac{1}{2} \frac{\mu}{1 - \mu} \frac{1 - \delta}{\delta} \tag{3.8}$$

Furthermore, we assume that

Only in state $S'' = S_c$ mobilization can affect the voting outcome, so we require that $0 < \beta < \alpha$.

This assumption means that in case of $S'' = S_a$ or $S'' = S_b$, the voting outcome is predetermined and mobilization has a meaning only in case of $S'' = S_c$. Note, that although it makes further analysis simple, this assumption doesn't affect the qualitative results.

Voting day. After candidates choose the levels of their mobilization spending the game turns to its passive phase. Turnout rates of core supporters, δ_A and δ_B, realize. The popularity shock switches to some state S'' from S'. The latent fraction s'' of swing voters supporting candidate A is realized conditional on the range defined by the state S'' and is getting revealed on the voting day. The candidate who gets more ballots wins and receives the payoff R. In case both candidates get equal amounts of votes, a fair coin decides the winner.

Timing. We summarize our basic model by providing the formal timing of the game. Game proceeds in two periods. The first period refers to the advertizing phase of the campaign and the second period relates to the voter mobilization. This sequence of actions is motivated by the timing of real electoral campaigns.[11]

---

[10]We note, that a realistic assumption would be to let mobilization success function q be depending on the natural turnout rate δ and the share of partizans μ, so $q(y; \delta, \mu)$. Intuitively, as long as natural turnout δ is higher, it should be more difficult and expensive for candidates to increase it further. This logic directly stems from the arguments described in the paragraph above for a strictly concave Φ, i.e. the marginal effect of y on $E\delta(y)$ drops as b (hence, δ) is getting larger. Furthermore, for a higher share of partizans, μ, candidates should spend more money to attain a given turnout rate (scale effect). However, we will not explicitly introduce these features since they don't change the qualitative implications of the model.

[11]We note, that one-shot game version of our model is possible, but requires some additional assumptions for the valid equilibrium characterizations.

1. [First period] In the beginning, candidates observe all exogenous parameters and distributions. Candidates simultaneously and independently raise and spend x_j in the advertizing stage. The purpose of this stage for candidates is to increase their constituencies among swing voters by persuasion (advertizing). Technically, candidates affect the split of swing voters' sympathies, s^0.

2. The latent split s^0 realizes and candidates receive a coarse signal S^0, which relates s^0 to a certain interval from the partition (3.2). Afterwards (or simultaneously), with probability γ the advantage factor shifts the latent split to the state S_b and candidates learn it in case it happens. The new state, S', is either stay S^0, or change to S_b. Note, that advertizing efforts of candidates directly affect the probability distribution of S' on the state-space (3.2).

3. [Second period] After candidates learn S' they simultaneously and independently raise and spend y_j for mobilization purposes. While mobilizing, candidates attempt to increase the probability of participation for their core supporters (partizans). However, for different realizations of swing voters' preferences at the voting day mobilization has distinct marginal effects on the probability of winning.

4. The Markov shock of popularity happens and the state of swing voters' support shifts from S' to S''. Conditional on being in state S'', the latent split s'' is determined. At the same time, the outcomes of candidates' mobilization campaigns realize. Citizens vote according to their sympathies and participation conditions. The winning candidate is revealed and candidates receive their payoffs.

3.3.3 Model Analysis

Although the model presented in the previous section may seem somewhat involved at first sight, it permits to readily derive several important qualitative implications. We start solving for an equilibrium from the second period, i.e. from the strategic choice of mobilization spending.

Mobilization. We note, that the realization of S' comprises all the relevant information that happens in the game up to the moment. Thus, strategies of candidates in the mobilization phase are state dependent and determined by a given state of relative popularity S'. Being in state S', candidate j solves the following problem

$$\max_{y_j} R\,G_j(y_j, y_{-j}|S') - cy_j \qquad (3.9)$$

where G_j denotes the probability of winning the elections for candidate j as a function of mobilization spending of candidates and the pre-mobilization state S'.

As we may remember, only state $S'' = S_c$ at the voting day implies the pivotal role of mobilization. Thus, mobilization can be treated as a state contingent security which pays only in case $S'' = S_c$. Then, candidate's mobilization investment y_j is a hedge against the state $S'' = S_c$ where an opponent may win elections by mobilizing her core supporters. Intuitively, if the probability of having $S'' = S_c$ is higher, the "demand" on mobilization is also higher and candidates invest more. The probability of having S_c at the voting day depends on the current state S' through the transition probability matrix P. Since an advertizing phase affects the realization of S' and the transition from S' to S'' has a memory due to assumption (3.3.2), it makes a bridge between advertizing and mobilization stages of the campaign.

It is a simple exercise to derive probability of winning G_j. For a given state S' we have

$$G_a(y_a, y_b|S') = \bar{e}'_{s'} P \bar{g} = \bar{e}'_{s'} P \begin{pmatrix} 0 \\ \frac{1}{2} + \frac{\beta}{2\alpha}\Big(q(y_a) - q(y_b)\Big) \\ 1 \end{pmatrix} \qquad (3.10)$$

where $\bar{e}_{s'}$ is 3×1 unit vector with the unit on the position of S' and with zeros otherwise. For candidate B we naturally have $G_b(y_a, y_b|S') = 1 - G_a(y_a, y_b|S')$. Vector $\bar{g}$ refers to the probabilities of winning conditional on the terminal popularity states S''.

For all states S' the probability $G_j(y_j, y_{-j}|S')$ is strictly concave in y_j, so the first order condition of candidate problem (3.9) possibly determines an equilibrium spending allocation. From equations (3.9) and (3.10) we may derive these conditions for different pre-mobilization states of relative support S':

$$\frac{\beta}{2\alpha} \frac{\partial q(y_j)}{\partial y_j} P_{S'|S_c} = \frac{c}{R}, \qquad S' \in \{S_b, S_c, S_a\} \qquad (3.11)$$

Equation (3.11) has at most one solution, y_j^*. It exists if ratio c/R is sufficiently small, so the office rent is much larger than marginal cost of raising a unit of campaign funds. Note, that it is usually the case in reality. In all possible states S' the first order conditions of candidates are symmetric due to our assumptions. Therefore, $y_j^* = y_{-j}^* = y^*$. For y^* to be a Nash equilibrium we should also require the robustness to deviations. Due to concavity of problem (3.9) the only possible profitable deviation is a corner strategy $y_j = 0$. Thus, we may formulate state dependent participation constraints for candidates

$$R\Big(G_j(y^*, y^*|S') - G_j(0, y^*|S')\Big) > cy^*, \qquad S' \in \{S_b, S_c, S_a\} \qquad (3.12)$$

The candidates may be idle in the mobilization phase depending on characteristics of electoral district. Note, that constraint (3.12) is more tight at states S_a and S_b where the probability of having a pivotal role of partizan participation is lower. Indeed, in our

empirical study we find that in about 30% of congressional electoral races from our sample candidates don't report any mobilization spending.

In case (3.12) holds, an interior allocation y^* (if exists) is Nash equilibrium. Equation (3.11) allows to derive several qualitative results for such case.

In case of interior equilibrium the characteristics of electoral environment influence the amount of mobilization spending (at intensive margin). In particular, an equilibrium mobilization spending y^* increases with the relative fraction of core supporters μ in the district and falls with the natural turnout rate δ. Also, y^* is higher in the competitive state $S' = S_c$, where relative popularity before the mobilization stage is not biased towards one of the candidates.

Advertizing. We now turn to the analysis of the advertizing phase of the campaign. In the beginning of the game candidates can perfectly foresee the state contingent expected utilities for all possible intermediate popularity states S'. The role of advertizing stage, therefore, is to increase the probability of having such S' which provides the highest expected utility. Naturally, for candidate A such desired state is $S' = S_a$ and for candidate B such state is $S' = S_b$.

Denote by $\bar{u}_j$ a 3×1 vector where each component is an expected utility conditional on getting into a given state S'. Using (3.4) we may write the problem which candidate solves at the advertizing stage

$$\max_{x_j} \bar{\pi}'_{S'} \bar{u}_j - c x_j \tag{3.13}$$

where $\bar{\pi}_{S'}$ stands for a discrete probability distribution of getting into state S' conditional on advertizing spending and also takes into account an advantage factor of candidate B. The first order condition of the candidate j problem can be easily derived and reads as

$$\varphi(1 - \gamma) \frac{\partial H(x_j, x_{-j}; \Delta)}{\partial x_j}(1 - 3\rho) = \frac{c}{R} \tag{3.14}$$

where $\Delta = (1-\mu)\delta$ is a mass of swing voters who expected to participate in voting. Similar to the mobilization case, if c/R is sufficiently small there exists a unique interior allocation $x^* = x_a^* = x_b^*$ which solves (3.14) for $j = \{A, B\}$. The problem (3.13) is concave in its argument, so x^* is Nash equilibrium in case candidates don't prefer a corner allocation $x_j = 0$ to the interior one $x_j = x^*$. The participation constraint is given by

$$\varphi(1 - \gamma) H(x^*, 0; \Delta) R(1 - 3\rho) > c x^* \tag{3.15}$$

Thus, as in the mobilization phase, we see that the strategic choice of advertizing spending at extensive margin heavily depends on the electoral environment. However, our empirical study shows that only a small fraction of candidates from our sample don't report any advertizing disbursement. It may suggest, for instance, that the real value of ρ is small

enough and the realization of S' almost certainly predicts S'' ("no surprise" environment).

If x^* exists and is Nash equilibrium, necessary condition (3.14) directly points out the following regularities in advertizing spending patterns

In case of interior equilibrium the characteristics of electoral environment influence the amount of advertizing spending (at intensive margin). In particular, an equilibrium advertizing spending x^* increases with the natural turnout rate δ and falls with the relative fraction of core supporters μ. Also, x^* is getting lower if advertizing technology is noisy (φ is lower) and the popularity shock is less persistent (ρ is higher).

Role of advantage. We now consider the implications of ex-ante advantage of one candidate over the other on the equilibrium campaign spending. First we note, that from (3.14) and (3.15) it is directly evident that an advantage parameter affects an advertizing decision of candidates both at extensive and intensive margins. Advertizing is higher if an advantage is lower, *ceteris paribus*. Intuitively, this is so because an ex-ante advantage of one candidate diminishes the marginal effect of advertizing spending on the probability distribution of relative popularity states S'. Therefore, candidates have less incentives to exert an advertizing effort.

A bit different logic is applied for mobilization. From (3.11) we know that equilibrium mobilization spending rises with $P_{S'|S_c}$. Thus, being at the states $S' = \{S_b, S_a\}$ candidates spend less than being at the state $S' = S_c$. Suppose now that candidate B has an advantage. It follows from (3.4) that on average the game comes at state $S' = S_b$ more often. Therefore, on average candidates spend less on mobilization if there is a clear leader in the race. We note that our model generates such effect not because an advantage directly affects an optimal choice of mobilization spending. An advantage shifts the distribution of swing voter preference realizations S' towards the leader and, as a result, an electoral race is more likely gets into the state where the marginal effect of mobilization activity is lower.

Advertizing spending are lower for electoral races with higher rates of ex-ante disparity (higher γ). It is also more likely, that mobilization spending are lower for such races.

We note, that these qualitative results may provide some theoretical grounds for our empirical evidence where such advantage effect is strongly supported. As a brief preview, our data says that in case of "safe" electoral races (such races, where there is a strong ex-ante leader) candidates spend significantly less on both advertizing and mobilization as compared to the cases of "close" races (such races, where there is no clear ex-ante favorite). We now turn to the empirical analysis where we investigate our qualitative findings from the modeling framework.

3.4   Empirical analysis

3.4.1   Working hypothesis, estimation methodology and data description

We now formulate two testable hypotheses based on the results from our analytical framework. Propositions 3.3.3 and 3.3.3 provide a link between candidates' campaign spending and several exogenous characteristics of an electoral district. In particular, these propositions relate both advertizing and mobilization disbursement with voter turnout and fractions of swing/partizan voters in a district. Due to the difficulties regarding the rigorous characterization of swing and partisan voters in the data, we suggest to gauge empirically the connection between campaign spending and voter turnout. In doing so we relate the natural turnout rate δ from the model with the expected turnout rate in a district.

Hypothesis 1. *Advertizing expenditures of candidates increase and mobilization expenditures decrease with expected turnout rate.*

Our second hypothesis is motivated by Proposition 3.3.3, which describes the relationship between the ex-ante competitiveness of the electoral race and campaign spending of candidates. We formulate the hypothesis in a following way:

Hypothesis 2. *Both advertizing and mobilization expenditures of candidates increase with expected competitiveness of the race.*

Candidate's expenditures on mobilization and candidate's spending on advertising are limited in their ranges: As already mentioned in section 3.2, in our sample ca. 27 percent of all candidates spend zero campaign funds on voter mobilization and roughly 5 percent of all candidates spend zero dollars on advertising. The choice of our estimation approach is driven by our theoretical framework, in which the candidates make a one-stage decision of how to allocate their available funds between mobilization and advertising under the non-negativity constraints for the two expenditures. In the model, some candidates would ideally spend negative amounts on one of the activities or both. We treat those candidates as candidates spending zero amounts in the data. This is why we use the Tobit regression model to test the two hypotheses above.[12]

We use the same data and the same selection criteria as we did in section 3.2 (appendix 3.6.2) to construct candidate's expenditures on mobilization (variable *mob*) and expenditures on advertising (*adv*). We divide the amount of disbursements on mobilization or advertising by the voting age population in a given congressional district in a given election cycle. This transformation makes the interpretation of expenditures more informative and intuitive.

---

[12]Our results persist if we use a two-stage Heckman correction model.

The natural turnout rate in the modeling framework implies a turnout rate in the absence of campaign effects. Since we can't directly observe it in the data, we use ex-post turnout rates from preceding House elections in the same congressional district (variable *past turnout*, where we take the turnout rates in 2008 for the 2012 election cycle and the turnout rates in 2006 – for the 2010 cycle to account for potential differences in voter turnout in midterm and presidential elections).

We cannot control explicitly for many factors that affect the ex-ante relative advantage in the model because we cannot observe them in the data (e.g. candidates' personal characteristics). To overcome this issue, we follow the strategy by Abramowitz (1991), who uses analysts' predictions from the *Congressional Quarterly* as a measure of expected advantage. However, this measure has two major problems. First, it is not truly endogenous to campaign spending, since the *Congressional Quarterly* predictions are made just a few weeks before the election and are affected by prior campaign expenditures. Second, this measure is narrowly defined because it only include three categories: *close*, *leaning* and *safe*. Instead, we use a *2010 and 2012 House Race Ratings* by *The Cook Political Report* as a proxy for the ex-ante degree of competitiveness.[13] To the best of our knowledge, we are the first to use this measure to control for expected electoral closeness.

The Cook Political Report is an independent, non-partisan organization that employs political analysts to monitor and analyze elections and campaigns for the US House of Representatives, US Senate, Governors and President. For the US House of Representatives elections, the Cook Political Report publishes district-by-district assessments of the status of the House races. During one year prior to the elections, the assessments are updated on a monthly basis. These assessments rate the state of each House race on a scale as being *Solid*, *Likely* or *Leaning* for either the Democrats (D) or Republicans (R), or as being a *Toss-up*. In a *solid* R/D race, the seats are not considered competitive. In a *likely* D/R race, the seats are not considered competitive, but have the potential to become engaged. A *leaning* D/R race is considered a competitive race, but one party has a slight advantage. Finally, a *toss-up* race is the most competitive race with either party having equal chances of winning.

Based on the four categories of the election assessments presented, we define seven candidate categories: *underdog*, *likely underdog*, *less likely underdog*, *close competitor*, *less likely leader, likely leader*, and *leader*. For example, if the race in a district is assessed to be *solid democratic*, then we define the candidate from the Democratic Party in this district as a *leader* and the candidate from the Republican Party – *underdog*. In a toss-up race, each candidate is defined as a *close competitor*.

These assessments have one important limitation in the context of our analysis: the *Cook Political Report* doesn't disclose the exact criteria on the basis of which the races

---

[13]See http://cookpolitical.com/house/charts/race-ratings/4402.

are assessed. Admittedly, the race assessment might take into account not only the factors, which make up our measure of ex-ante relative advantage in the model but some other factors. Some of these factors might not overlap with the explanatory variables that we will include in the regression. To the extent, that the analysts at the *Cook Political Report* observe the same factors that candidates in our sample observe when they make their spending decisions, this is not problematic for our analysis. Some of the factors might well overlap with our explanatory variables. This might, of course, induce multicollinearity between the race assessments and other covariates and result in imprecisely measured parameter estimates. In appendix 3.6.4, we specify a multi-logit model to regress the assessments on potentially influential factors, for which the data is available. One conclusion from this exercise is that available explanatory variables explain only a small portion of variation in the assessments.

During an election campaign, the assessments are themselves affected by the amounts spent on advertising and mobilization by the two candidates. To overcome potential endogeneity problems in our analysis, we use the assessments dated January 15th, 2010 and January 13th, 2012, roughly ten months before the House elections (the House elections of 2010 and 2012 were held on November 2nd, 2010 and November 6th, 2012, respectively). We prefer not to take earlier assessments because they neglect reapportionment and redistricting changes in some states. There is a small fraction of races, for which no assessment is available for the month January. For these races, we take the assessments dated June 4, 2010 and June 7, 2012, respectively.

The summary statistics for disbursements on voter mobilization and advertising for different categories of candidates are shown in table 3.1 and 3.2, respectively.

| | House 2010 | | | | House 2012 | | | |
| | Positive spending | | | Zero | Positive spending | | | Zero |
	Mean	Median	N	N	Mean	Median	N	N
Underdog	3.1	1.8	107	112	3.8	1.2	106	111
Likely underdog	9.6	7.5	51	11	8.1	7.0	49	9
Less likely underdog	16.3	11.7	29	3	12.6	11.3	25	4
Close competitor	16.1	16.9	31	3	21.6	20.7	42	4
Less likely leader	22.4	20.5	32	1	17.3	14.3	28	2
Likely leader	23.4	22.0	59	5	12.2	9.6	54	5
Leader	8.7	5.8	217	67	8.8	5.3	227	69
All	11.0	7.3	526	202	9.7	6.3	531	204

Table 3.1: Disbursements of candidates on voter mobilization (cents per voting age population), by candidate category
Mean and *Median* amounts are given in 2012 U.S. cents per district's voting age population. *Positive spending* refers to the candidates, who spend a positive amount on voter mobilization; *Zero* refers to those, who spend zero. *N* – number of candidates.

| | House 2010 | | | | House 2012 | | | |
| | Positive spending | | | Zero | Positive spending | | | Zero |
	Mean	Median	N	N	Mean	Median	N	N
Underdog	18.8	4.9	202	17	15.6	2.3	191	26
Likely underdog	126.2	108.5	61	1	86.9	74.0	56	2
Less likely underdog	178.1	169.7	32	0	153.1	139.9	29	0
Close competitor	209.1	199.0	34	0	251.3	234.4	46	0
Less likely leader	235.6	238.9	33	0	210.6	193.8	30	0
Likely leader	255.8	239.6	63	1	133.5	125.6	58	1
Leader	35.5	13.9	277	7	42.3	16.9	285	11
Total	82.7	26.4	702	26	71.9	17.5	695	40

Table 3.2: Disbursements of candidates on advertising (cents per voting age population), by candidate category
Mean and *Median* amounts are given in 2012 U.S. cents per district's voting age population. *Positive spending* refers to the candidates, who spend a positive amount on advertising; *Zero* refers to those, who spend zero. N – number of candidates.

Several interesting findings emerge. One can see from table 3.1 that the fraction of candidates who spent zero dollars on mobilization and those who spent a positive amount was roughly the same in the two election cycles – 40 and 60 percent, respectively. Conditional on mobilizing, spending was on average 11 cents per voting age individual in the mid-term 2010 election. Once we disaggregate expenditures by candidate category, we notice that there is significant heterogeneity in spending, which ranges from 3.1 cents to 23.4 cents. Observe that underdogs and leaders spend least amounts (3.1 and 8.7 cents, respectively), while those who are likely to lead the race spend the most (23.4 cent). Average spending was slightly lower in the 2012 presidential election (9.7 cents). As with the 2010 election, disbursements among underdogs and leaders were the lowest, while close competitors on average spent the highest amounts on voter mobilization. Table 3.2 reveals significant differences in expenditures on advertising as opposed to mobilization. Only a fairly small fraction of candidates spent nothing on advertising in two election cycles (roughly 5 percent). The average spending for those who advertised was as much as 82.7 cents in 2010, which is 7.5 times larger than for mobilization. As with mobilization expenditures, underdogs and leaders spent least, while likely leaders (for 2010 election) and close competitors (2012 election) spent most on advertising. Summarizing, tables 3.1-3.2 suggest that expenditures on both advertising and mobilization are lower in electoral races, in which one candidate has a clear advantage. This holds for the two election cycles. In 2012, we also observe that spending for both mobilization and advertising is the highest in toss-up elections. This speaks in favor of our hypotheses 1 and 2. Below we run a Tobit regression model to test these and other hypotheses formally.

It can be seen from tables 3.1-3.2 that *less likely underdogs* and *less likely leaders* form a relatively small (in terms of number of observations) group of candidates. For

example, there are overall 219 *underdogs* during the 2010 House elections in the sample, while there are only 32 *less likely underdogs*, 32 *close competitors* and 33 *less likely leaders*. For the categories to be of roughly equal size, we pool together the observations from the categories *less likely underdogs*, *less likely leaders* and *close competitors*. We set this newly built category to be the reference category in the regression analysis below and generate dummy variables for the remaining categories (*underdog*, *likely underdog*, *leader*, and *likely leader*).

We use a broad set of other control variables.[14] We control for the advantages that a candidate might enjoy on account of her status as an incumbent running for re-election (dummy variable *incumbent*) and for candidate's party affiliation (dummy variable *democrat*). We also need to control for the total campaign budget of a candidate, since a candidate with a larger budget will tend to spend more on both advertising and mobilization. As a proxy for the campaign budget of a candidate, we use candidate's total administrative, salary and overhead expenses (variable *overhead expenditures*). In our data set, such expenses typically include rent, staff salaries, office supplies, equipment, furniture, legal and accounting expenses. Since a portion of these expenses might be endogenous to advertising and mobilization expenditures, we drop those overhead expenses that occur after June 1st of the year, in which the election takes place.

We introduce a dummy variable *presidential election* (which takes value of 1 for the 2012 election cycle), since the spending pattern is likely to change in anticipation of spending activities of the presidential candidates. We also include regional dummies according to the Census classification (*North-East*, *Mid-West*, *South* and *West*). Since our model specification doesn't contain a constant, we include all four regional dummies, as well as interaction terms between regional dummies and the dummy variable *democrat* (with the dummy variable *Democrat–Mid-West* being the reference category). Furthermore, we use a set of variables, which describe the economic and the demographic conditions at the level of congressional districts. These variables affect voters' turnout rates, voter's responsiveness to advertising and mobilization as well as voter's preferences towards the candidates.[15]

3.4.2   Empirical results

Our empirical results are shown in table 3.3. We regress mobilization expenditures on a set of described explanatory variables using the Tobit model. This specification is denoted

---

[14]The sources for the data are described in appendix 3.6.3.

[15]We include unemployment rate (*unemployment*), share of people of age 25-64 with at least some college education in the total population of age 25-64 (*education*), median household income in 2012 adjusted dollars (*median income*), shares of hispanic and black population of age above 18 in the total population of age above 18 (the variables *hispanic* and *black*, respectively), share of population above age 60 in the total population of age above 18 (*elderly*).

mob (Tobit) in the table. We also run the OLS model, while we restrict the sample
to include only those candidates who spend positive amounts on mobilization.   This
specification is denoted *mob (OLS)*. We repeat the analysis for advertising expenditures.
These specifications are labelled *adv (Tobit)* and *adv (OLS)*, respectively. Recall that the
category of *toss-up* races was defined to be a reference category in all four regressions.

We find that in both Tobit and OLS specifications the voter turnout in previous
elections has a significant positive effect on advertizing expenditures, which is consistent
with Hypothesis 1. However, we don't find any significant effect of past voter turnout on
mobilization disbursements.

Consistent with our Hypotheses 2, we find that on average spending on each form
of campaign expenditures rises as races become more competitive. More specifically,
we find that safe underdogs and safe leaders spend ceteris paribus least amounts on
advertizing and mobilization (the corresponding estimates are significant at 1 percent
level in all 4 specifications). Likely underdogs spend higher amounts on both advertizing
and mobilization than safe underdogs but lower amounts than close competitors (the
corresponding estimates are significant at 1 percent level in all 4 specifications).

3.5   Outlook

The modeling framework that we develop in this paper may be extended in several di-
rections to analyze other important questions. Firstly, the convenient mechanics and
reasonable intuition from our simple model can be used to build a more general frame-
work of a multi-period campaign spending game. A generic number of popularity states
and a certain number of rounds can be introduced there, so in each period candidates
observe the realizations of Markov switching states and revise their campaign strategies
in response to the realized stochastic shocks and expected transitions of the relative pop-
ularity. Such a generalization of our basic framework may be useful for the analysis of
dynamic properties of advertizing and/or mobilization spending strategies of candidates
in the course of the long-run electoral races (like Presidential campaigns).

Another interesting direction is to consider voters as rational agents who infer noisy
signals about the current states of the electoral balance in the race and then make the
decisions about campaign contributions in the first stage and participation in voting in the
last stage. Similar to candidates in our present model, voters assess the probabilities to
affect the voting outcome by doing these actions (contributions and voting) and derive the
corresponding marginal effects on the probability of winning for the preferred candidate.
This approach may be helpful for endogenizing the campaign contributions (supply side
of the candidates' fundraising) as well as the voting participation of citizens, so it directly
pretends to unveil the celebrated "paradox of voting". Also, this extended framework may

	mob (Tobit)	mob (OLS)	adv (Tobit)	adv (OLS)
Degree of competition:				
safe underdog	-0.184***	-0.132***	-1.803***	-1.786***
	(0.000)	(0.000)	(0.000)	(0.000)
likely underdog	-0.096***	-0.089***	-1.004***	-0.999***
	(0.000)	(0.000)	(0.000)	(0.000)
likely leader	-0.028	-0.032*	-0.543**	-0.545**
	(0.143)	(0.075)	(0.023)	(0.024)
safe leader	-0.127***	-0.111***	-1.981***	-1.971***
	(0.000)	(0.000)	(0.000)	(0.000)
Characteristics of candidate:				
incumbent	0.003	0.004	0.370***	0.375***
	(0.825)	(0.752)	(0.008)	(0.008)
Democrat	0.019	0.064***	0.009	-0.034
	(0.115)	(0.000)	(0.924)	(0.756)
overhead expenditures	0.021***	0.037***	0.089	0.094
	(0.005)	(0.000)	(0.131)	(0.121)
Characteristics of elections:				
presidential election	-0.014	-0.009	-0.279***	-0.254***
	(0.246)	(0.436)	(0.003)	(0.008)
past turnout	0.001	0.000	0.014***	0.014***
	(0.102)	(0.634)	(0.000)	(0.000)
Observations	1019	747	1019	1012
R^2		0.6433		0.5832
F	26.803	52.433	41.317	40.870

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Omitted groups: regional dummies, demographic and economic characteristics of districts, interaction terms.

Table 3.3: Estimation results for alternative model specifications

shed an additional light on the equilibrium campaign strategies of candidates.

Finally, our framework (probably extended along the lines outlined above) can become a useful tool for a normative analysis of campaign spending reforms. We relate these interesting research directions for the future work.

3.6 Appendix

3.6.1 Voter turnout rates

To generate figure 3.1, we used two separate data sets. The first data set is *Federal Elections 2012: Election Results for the U.S. President, the U.S. Senate, and the U.S. House of Representatives* by the Federal Election Commission (http://www.fec.gov/pubrec/ fe2012/federalelections2012.shtml). This data set contains the information on the total number of votes collected by each candidate during the general election to the House of Representatives in 2012. We compute the total number of casted ballots (including scattered ballots) by each congressional district. The second data set contains voting age population by congressional district. The data come from the American Community Survey (http://www.census.gov/). The data for each year are contained in the Table *B05003. Sex by age by nativity and citizenship status*, which is part of 1-year estimates data sets. To obtain the voting age population in a congressional district, we sum the following variables: *Native males, 18 years and over (HD01_VD09), Foreign born males, 18 years and over, naturalized U.S. citizens (HD01_VD11), Native females, 18 years and over (HD01_VD20), Foreign born females, 18 years and over, naturalized U.S. citizens (HD01_VD22)*. We define the voter turnout rate as the ratio of the total number of casted ballots to the voting age population.

3.6.2 Candidates' disbursements

The data on candidates' disbursements during House of Representatives elections are publicly accessible through the homepage of the Federal Election Commission under http://www.fec.gov/data/DataCatalog.do?cf=downloadable (Campaign Finance Disclosure Portal → Downloadable Data → Data Catalog → Candidate Disbursements).

The election cycle of 2010 is the first, for which the Federal Election Commission files the disbursement data electronically. We include the 2010 and 2012 election cycles into the analysis, because there is still missing data for the later election cycles at the time this draft of the paper is completed. We exclude the Insular areas (Guam, Northern Mariana Islands, and Virgin islands) and Washington D.C. We keep only those candidates in the sample, who won the primary elections (if such took place). However, when computing the

disbursements of a candidate, we do not disregard the expenditures that this candidate incurred during the primary elections (if such took place). We drop those candidates from the sample, who are not members of the Republican or the Democratic Party. We also exclude the races, in which two candidates from the same party ran for the office. Our sample contains the incumbent/challenger races, as well as unopposed elections or elections without incumbent.

Our final sample includes 1,463 observations. Roughly 85 percent of these observations correspond to the candidates, who run in a *classic* race with a challenger and an incumbent. Ca. 13 percent of the candidates compete in elections without incumbent (because the incumbent retired, died, or resigned or because a new congressional district is created). The remaining 2 percent of the candidates have no opponent.

Each candidate's expenditure is assigned to one of 12 categories: administrative/salary/overhead expenses, advertising expenses, polling expenses, travel expenses, solicitation and fundraising expenses, campaign materials, campaign event expenses, donations, transfers, loan repayments, refunds of contributions, and uncategorized disbursements. We assign all items of the category *polling expenses* to expenditures on voter mobilization. In our data set, most expenditures of this category are on door-to-door canvassing, voter filing, voter research, phone banks, and get-out-the-vote activities. We assign all items of the category *advertising expenses* to the expenditures on advertising. Besides the category, each expenditure contains a purpose. We add to mobilization expenditures those uncategorized disbursements, whose reported purpose includes one of the following key words: canvassing, voter mobilization, grass rooting, getting out the vote, guiding and walking voters, conducting voter surveys. We add to advertising expenditures all those uncategorized expenditures, whose reported purpose includes one of the following the key words: advertising, TV ad, radio ad, media purchase. We drop those expenditures, which have a missing candidate ID. All expenditure amounts are converted to 2012 U.S. dollars using the Consumer Price Index.

3.6.3 Control variables

The incumbency status and party affiliation of a candidate are contained in the expenditure data by the Federal Election Commission (see above). To construct the voter turnout rates in 2006 and 2008 we proceed the same way as we did to plot figure 3.1.

The variables, which describe the demographic and economic conditions at the level of congressional districts, were constructed using the one year estimates by the American Community Survey (http://www.census.gov/). The data on the unemployment rate, the educational level, and the share of hispanic population were taken from Table *S2301. Employment status.* Table *S0102. Population 60 years and over in the United States*

contains the data on the share of old population. The data on the white, black and asian population were taken from Table *B02001. Race.* Table *B19013. Median household income in the past 12 months* contains the data on the median household income, where we convert the dollar amounts for the year 2010 to 2012 using the Consumer Price Index.

The Census regions (variables *North-East, Mid-West, South, West*) were obtained from the Census Bureau Regions and Divisions at http://www2.census.gov/geo/docs/ maps-data/maps/reg_div.txt.

3.6.4   Cook Political Report Ratings

	competitiveness
Characteristics of candidate:	
incumbent	-1.100***
	(0.000)
Democrat	-0.032
	(0.902)
overhead expenditures	1.109***
	(0.000)
Characteristics of elections:	
presidential election	-0.719***
	(0.009)
past turnout	0.025**
	(0.032)
Observations	1019
Pseudo R^2	0.1134

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Variable *competition* has three realizations: 1) *lowest* (*safe underdog* and *safe leader*), 2) *medium* (*likely underdog, likely leader*) and 3) *highest* (*toss-up, less likely underdog, less likely leader*). Omitted groups: regional dummies, demographic and economic characteristics of districts, interaction terms.

Table 3.4: Ordered multi-logit regression

Bibliography

Abramowitz, A. I. Incumbency, campaign spending, and the decline of competition in u.s. house elections. *The Journal of Politics*, 53(1):34–56, 1991.

Aiyagari, S. R. Uninsured Idiosyncratic Risk and Aggregate Saving. *QJE*, 109(3):659–84, 1994.

Azzimonti, M., de Francisco, E., and Krusell, P. Median-voter Equilibria in the Neoclassical Growth Model under Aggregation. *Scandinavian Journal of Economics*, 108(4): 587–606, 2006.

Bachmann, R. and Bai, J. Politico-Economic Inequality and the Comovement of Government Purchases. *Review of Economic Dynamics*, 16:565–580, 2013.

Bartels, L. *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton University Press, 2009.

Benabou, R. Unequal Societies: Income Distribution and the Social Contract. *American Economic Review*, 90(1):96–129, 2000.

Benabou, R. and Ok, E. Social Mobility And The Demand For Redistribution: The Poum Hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487, 2001.

Bourguignon, F. and Verdier, T. Oligarchy, democracy, inequality and growth. *Journal of Development Economics*, 62(2):285–313, 2000.

Campante, F. R. Redistribution in a model of voting and campaign contributions. *Journal of Public Economics*, 95:646 – 656, 2011.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J.-V. Accounting for the U.S. Earnings and Wealth Inequality. *Journal of Political Economy*, 111(4):818–857, 2003.

Conesa, J. C. and Krueger, D. Social Security Reform with Heterogeneous Agents. *Review of Economic Dynamics*, 2(4):757–795, 1999.

Conesa, J. C. and Krueger, D. On the optimal progressivity of the income tax code. *Journal of Monetary Economics*, 53(7):1425–1450, 2006.

Cooley, T. and Prescott, E. Economic growth and business cycle. In *Frontiers of business cycle research*. Princeton University Press, 1995.

Corbae, D., D'Erasmo, P., and Kuruscu, B. Politico-economic consequences of rising wage inequality. *Journal of Monetary Economics*, 56(1):43–61, 2009.

Dahl, R. *Who Governs? Democracy and Power in an American City*. Yale University Press, 1961.

Diaz, A. and Luengo-Prado, M. J. The Wealth Distribution With Durable Goods. *International Economic Review*, 51(1):143–170, 2010.

Domeij, D. and Heathcote, J. On The Distributional Effects Of Reducing Capital Taxes. *International Economic Review*, 45(2):523–554, 2004.

Elo, I. T. and Preston, S. H. Educational differentials in mortality. *Social Science & Medicine*, 42(1):47–57, 1996.

Erikson, R. and Palfrey, T. Equilibria in campaign spending games: Theory and data. *The American Political Science Review*, 94(3):595–609, 2000.

Gonzalez-Eiras, M. and Niepelt, D. The future of social security. *Journal of Monetary Economics*, 55(2):197–218, 2008.

Gouveia, M. and Strauss, R. Effective federal individual income tax functions: An exploratory empirical analysis. *National Tax Journal*, 47(2):317–39, 1994.

Grossman, G. M. and Helpman, E. Electoral competition and special interest politics. *The Review of Economic Studies*, 63(2):265–286, 1996.

Guvenen, F. An empirical investigation of labor income processes. *Review of Economic Dynamics*, 12(1):58–79, 2009.

Hansen, G. D. The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked. *Journal of Applied Econometrics*, 8(1):71–80, 1993.

Hassler, J., Storesletten, K., and Zilibotti, F. Democratic public good provision. *Journal of Economic Theory*, 133(1):127–151, 2007.

Heathcote, J., Perri, F., and Violante, G. L. Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006. *Review of Economic Dynamics*, 13(1):15–51, 2010a.

Heathcote, J., Storesletten, K., and Violante, G. L. The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4): 681–722, 2010b.

Herrera, H., Levine, D. K., and Martinelli, C. Policy platforms, campaign spending and voter participation. *Journal of Public Economics*, 92(3-4):501–513, 2008.

Herrera, H., Morelli, M., and Palfrey, T. Turnout and Power Sharing. *Economic Journal*, 124(574):F131–F162, 2014.

Huggett, M. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969, 1993.

Huggett, M. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, 1996.

Klein, P., Krusell, P., and Rios-Rull, J.-V. Time-Consistent Public Policy. *Review of Economic Studies*, 75(3):789–808, 2008.

Kopczuk, W., Saez, E., and Song, J. Earnings Inequality and Mobility in the United States: Evidence from Social Security Data Since 1937. *The Quarterly Journal of Economics*, 125(1):91–128, 2010.

Krusell, P. and Rios-Rull, J.-V. On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model. *American Economic Review*, 89(5):1156–1181, 1999.

Lagunoff, R. and Bai, J. Revealed Political Power. *International Economic Review*, 54: 1085–1115, 2013.

Lindbeck, A. and Weibull, J. Balanced-budget redistribution as the outcome of political competition. *Public Choice*, 52(3):273–297, 1987.

Lindbeck, A. and Weibull, J. W. A model of political equilibrium in a representative democracy. *Journal of Public Economics*, 51(2):195–209, 1993.

Meirowitz, A. Electoral Contests. Papers 06-21-2007, Princeton University, Research Program in Political Economy, 2006.

Meltzer, A. and Richard, S. Rational Theory of Size of Government. *Journal of Political Economy*, 89(5):914–27, 1981.

Mueller, D. *Public Choice III*. University Cambridge Press, 2003.

Nalebuff, B. and Shachar, R. Follow the Leader: Theory and Evidence on Political Participation. *American Economic Review*, 89(3):525–547, 1999.

Page, B., Bartels, L., and Seawright, J. Democracy and the Policy Preferences of Wealthy Americans. *Perspectives on Politics*, 11:51–73, 2013.

Quadrini, V. Comment on: "Politico economic consequences of raising wage inequality" by Corbae, D'Erasmo and Kuruscu. *Journal of Monetary Economics*, 56(1):62–65, 2009.

Quadrini, V., Rios-Rull, J.-V., and Diaz-Gimenez, J. Dimensions of inequality: Facts on the US distributions of earnings, income, and wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2):3–21, 1997.

R. Glenn Hubbard, Jonathan Skinner, S. P. Z. Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):360–399, 1995.

Rios-Rull, J.-V. and Krusell, P. On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model. *American Economic Review*, 89(5):1156–1181, 1999.

Rosen, S. Prizes and incentives in elimination tournaments. *American Economic Review*, 76:701–715, 1986.

Rosenstone, S. and Hansen, J. *Mobilization, Participation, and Democracy in America.* Macmillan, New York, 1993.

Schultz, C. Strategic Campaigns and Redistributive Politics. *Economic Journal*, 117 (522):936–963, 2007.

Skaperdas, S. and Grofman, B. Modeling negative campaigning. *American Political Science Review*, 89:49–61, 1995.

Snyder, J. Election goals and the allocation of campaign resources. *Econometrica*, 57: 637–660, 1989.

Song, Z., Storesletten, K., and Zilibotti, F. Rotten Parents And Disciplined Children: A Politico-Economic Theory Of Public Expenditure And Debt. *Econometrica*, 80(6): 2785–2803, 2012.

Spenkuch, J. and Toniatti, D. Political Advertising and Election Outcomes: Evidence from Thousands of Natural Experiments. *Working Paper*, 2015.

Tauchen, G. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181, 1986.

Tullock, G. Efficient rent seeking. *In James M. Buchanan, Robert D. Tollison, Gordon Tullock, (Eds.), Toward a theory of the rent-seeking society. College Station, TX: Texas A & M University Press*, 57:97–112, 1980.

Wolff, E. Changes in Household Wealth in the 1980s and 1990s in the U.S. Economics Working Paper Archive wp_407, Levy Economics Institute, The, 2004.