

The Politics of Flat Taxes*

Daniel R. Carroll[†]

Jim Dolmas[‡]

Federal Reserve Bank of Cleveland

Federal Reserve Bank of Dallas

Eric R. Young[§]

University of Virginia and Zhejiang University

August 25, 2017

Abstract

We study the determination of flat tax systems using a workhorse macroeconomic model of inequality. Our first result is that, despite the multidimensional policy space, equilibrium policies are typically unique (up to a fine grid numerical approximation). The majority voting outcome features (i) zero labor income taxation, (ii) simultaneous use of capital income and consumption taxation, and (iii) generally low transfers. We discuss the role of three factors – the initial heterogeneity in sources of income, the mobility of income and wealth, and the forward-looking aspect of voting – in determining the equilibrium mix of taxes.

*We thank seminar participants at many places for comments. Young thanks the Bankard Fund for Political Economy for financial support. This paper does not represent the views of the Federal Reserve Bank of Cleveland, the Federal Reserve Bank of Dallas, or the Federal Reserve System.

[†]Email: daniel.carroll@clev.frb.org.

[‡]Email: jim.dolmas@dal.frb.org.

[§]Email: ey2d@virginia.edu.

1 Introduction

Don't tax you, don't tax me, tax that fellow behind the tree." - Russell B. Long

Tax systems across the world are quite diverse, not only in terms of the overall level of tax rates but also in terms of the mix of tax systems. Carey and Rabesona (2002) document a range of tax systems from OECD countries, which we plot here in Figure (1); tax systems range from low (South Korea) to high (Norway) in terms of overall levels, and also vary substantially in terms of the mix between consumption, labor income, and capital income rates. The range of tax rates are $[6.4, 25.7]$ for consumption, $[9.9, 49.6]$ for labor income, and $[12.9, 39.5]$ for capital income. There is little correlation between the rates within a country as well. Similar diversity appears in the data constructed by Mendoza, Razin, and Tesar (1994) and Trabandt and Uhlig (2011).

A natural question to ask is how these tax systems arose. Capital income taxes, for example, are often viewed as inefficient (Atkeson, Chari, and Kehoe 1999), while consumption taxes are generally viewed as a more efficient revenue source.¹ One is then led to ask why, if capital income taxes are inefficient, do some countries nevertheless impose high rates? And, conversely, why are consumption taxes not used more intensively if they are superior? And what about labor income taxes? And finally, why are the taxes so different across countries? We go looking for the answers to these questions through voting, where questions of efficiency are placed to the side in favor of political feasibility.

The workhorse model of inequality is Aiyagari (1994), which we extend to include an array of fiscal instruments and elastic labor supply. We apply solution concepts for majority voting in multiple dimensions to study what kinds of tax systems could arise in a once-and-for-all vote, with voters taking into account the transitional and long-run consequences of particular tax systems. In a one-dimensional policy space, we could employ

¹The capital tax inefficiency result does not hold in the presence of incomplete insurance and/or redistributive motives (see Golosov, Kocherlakota, and Tsyvinski 2003, Chien and Lee 2006, Conesa, Kitao, and Krueger 2008, and Straub and Werning 2014). Results on the efficiency of consumption taxation can be found in Krusell, Quadrini, and Ríos-Rull (1996), Correia (2010), and Chari, Nicolini, and Teles (2016).

the median voter theorem to obtain a Condorcet winner (provided preferences are single-peaked). Unfortunately, the existence of a Condorcet winner in more than one dimension is guaranteed only under extreme symmetry conditions (called Plott symmetry after Plott 1967), and in our model preferences over taxes are derived from primitive preferences over consumption and leisure and cannot simply be assumed to possess the necessary symmetry.² Moreover, we cannot appeal to a single-crossing property because households in our model are heterogeneous along two dimensions (labor productivity and wealth) and cannot be ordered according to their preferences over taxes. For example, a household with high wealth and low productivity views the tradeoff between capital and labor income taxes differently than does a household with high productivity and low wealth, even if the two households have the same current resources. If a Condorcet winner does not exist, we want to find all the political equilibria. We therefore turn to generalized solution concepts of majority voting that permit multiple winners (that is, they are set-valued).

We use the most restrictive of the solution concepts in the literature, the essential set from Dutta and Laslier (1999). The essential set is given by the strictly positive support of strategies played in any mixed strategy Nash equilibrium of the voting game (a zero-sum game in which the payoff is one if your preferred strategy wins and minus one if it loses). This set is contained within the uncovered set from McKelvey (1986), which contains the strategies that would survive under agenda setting and sophisticated voting (that is, what strategies could be implemented if we selected different orderings for pairwise competition), and so has a convenient interpretation as those strategies that could win depending on which agent controls the voting agenda.³ If the uncovered set is a singleton (and therefore so is the essential set), then the unique element is the Condorcet winner.⁴

²We do not attempt to find sufficient conditions for Plott symmetry in our model. Given that we can only solve the model numerically, it seems impossible to make any progress along this line.

³Because the uncovered set allows for multiple solutions, we can sidestep the thorny problem of how to decide who gets to decide the agenda, which obviously leads to an infinite regress problem.

⁴Miller (2007) contains an extensive discussion of the geometry of the uncovered set. He also defines the notions of "locally covered" and "covered at a distance" that influence the shape of the set. Given that our uncovered set is a singleton, these considerations have not arisen.

We first list our three main results, then discuss the features of the model that are crucial for obtaining them. First, we find a Condorcet winner – at least up to a fairly fine grid approximation, outcomes are unique, and any variation we miss due to the discrete approximation is small enough that it has no economic consequences. Second, the Condorcet winner chooses to set labor income taxes to zero, relying on a combination of capital income and consumption taxes. And third, the Condorcet winner sets lump-sum transfers to zero in the long run, after one period of positive transfers. These results arise as a combination of three factors – the shape of the initial distribution (in terms of income source and the marginal propensity to consume), the mobility of earnings and wealth of agents (how quickly the poor become the middle class), and the patience of agents (how forward-looking their voting behavior is).

Why do we obtain a Condorcet winner? In a model with a small number of types, one can in principle assemble a majority in many different ways. If those majorities can be assembled using types with very different preferred tax systems, then the range of equilibrium outcomes can be large. Thus, an agenda setter, by properly sequencing votes, can obtain very different equilibrium policies. Our model endogenously delivers a large number of types, but the measure of these types is such that effective majorities cannot be constructed that prefer very different outcomes – the winning coalition must include the middle class, who in general prefer the same kinds of taxes as the poor. The rich are quite different but insufficiently numerous to win, even though our model somewhat understates the measure of rich agents.⁵

Why does the Condorcet winner have zero labor taxes? Here, patience and mobility play an important role. Our vote is once-and-for-all, meaning that once in place the taxes cannot be changed. Even without commitment, the currently "middle class" are reluctant to tax labor income since that is their primary source of income (labor's share of total

⁵In the appendix we show that it does not require very many types for the range of outcomes to be heavily restricted; specifically, going from four types to nine types in Dolmas (2008) is sufficient to get (nearly) a Condorcet winner.

income in our model is 0.64, and is higher for most agents and lower for the very wealthy). With commitment, even the poor become reluctant to tax labor income, as they anticipate their wages to rise (due to mean reversion in idiosyncratic productivity) and they place relatively high weight on future periods (due to the low discounting needed to match the aggregate wealth-income ratio). Only the currently wealthy wish to impose labor taxes as they do not work, but these agents are not sufficiently numerous to win. Furthermore, even the non-working wealthy do not like very high labor taxes, because these taxes reduce the return to capital.

Why does the Condorcet winner choose both capital income and consumption taxes? Both taxes attack the inelastic initial distribution of capital, meaning that they are attractive to the initially poor. The consumption tax, however, is regressive – it affects the initially poor more than the initially wealthy because the marginal propensity to consume is high for the poor. As a result, voters are reluctant to impose high consumption taxes. In contrast, the capital income tax is progressive, but has negative long-run effects by reducing wages; as with labor income taxation, low wages are opposed by most voters. The result is a compromise that is reminiscent of the theory of the second best (Lipsey and Lancaster 1956), where all instruments are used a little rather than one instrument a lot.

Why does the Condorcet winner choose zero transfers in the long run? Again, patience and mobility are crucial. Similar to labor taxes, transfers are limited by the expectation of the poor that they will be middle class in the near future. The middle class pays more in taxes than they receive in transfers (in general), and therefore oppose a transfer; many of the poor agree since they expect to be middle class soon. The wealthy are, of course, also opposed, since they will pay more in taxes than they receive in transfers. Note that zero transfers are chosen despite the presence of substantial market incompleteness and wage risk; as a result, the distribution of wealth changes very little over the transition.⁶

We conduct experiments designed to illuminate results two, three, and four. First, we

⁶Davila et al. (2012) shows that there are significant welfare gains available to a planner that cannot transfer resources between agents.

examine a version of the model with wage risk shut down – there is still heterogeneity, but it is now time-invariant.⁷ As shown in Chatterjee (1994) and Krusell and Ríos-Rull (1999), the resulting model has no mobility; if one household is initially wealthier than another, they remain so forever.⁸ In this case, we find that transfers increase substantially, with the concomitant increase in capital income and consumption taxes, but labor income taxes remain zero. Furthermore, if we impose that income taxation must be uniform – the only available choices tax capital and labor at equal rates – we get complete dependence on consumption taxation. Finally, we study a version where agents vote "myopically"; their economic decisions are made using one discount factor, while their votes are evaluated using a lower one. The result is that transfers again rise, but labor income taxes are still zero. Thus, we conclude that the zero labor tax result is very robust and is driven by the shape of the initial distribution, while the zero transfer result is more fragile and depends critically on mobility and patience.

With this result in mind, we turn to confronting our model with mobility facts. Using the PSID we estimate mobility for both earnings and wealth across quintiles over five-year horizons (as in Budria Rodriguez et al. 2002). Our model overstates slightly the amount of mobility over five-year horizons, but the support for the Condorcet winner is strong enough that a closer fit would be unlikely to change the result. Furthermore, the quintiles for which the deviations are largest are the top and bottom; under equal-weighted voting these groups are not decisive.⁹

We then study how outcomes would differ under "wealth-weighted" voting; as in Bachmann and Bai (2013), we assume that wealthier agents effectively have "more votes" than poorer agents. Here, we find that labor income taxes win – as votes become more concen-

⁷Note that this setup is not quite equivalent to complete markets, as under complete markets leisure would fluctuate with wages but the agent would be compensated with consumption via contingent claims. Since labor supply is convex, Jensen's inequality implies some small differences emerge in total labor input.

⁸This result is not general; see discussions in Caselli and Ventura (2000) or Carroll and Young (2011).

⁹Carroll and Young (2016) explore more in depth the fit between the basic Bewley model and the mobility data.

trated in wealthy agents, consumption and capital income taxes are set to zero and labor income taxes are used to finance government spending. Transfers remain zero. This result cements our intuition that the initial distribution is key for the zero labor tax result.

Finally, we consider an alternative environment suitable for multidimensional policy spaces, probabilistic voting. Under probabilistic voting, agents receive non-economic shocks to the utility of each policy vector, and the goal of candidates is to maximize the expected vote share their proposal receives. We explore this setup for two reasons: (i) many other papers use this structure, so it is important for connecting our work to the broader political economy literature, and (ii) the outcomes of probabilistic voting are equivalent to those obtained by maximizing a social welfare function, giving our interpretation of the equilibrium mix of capital and consumption taxation normative importance as a second-best intervention – voters are equalizing the net marginal gains from the two types of taxes, which occurs at positive values for both. We consider two cases, additive and multiplicative non-economic shocks, and find that both cases lead to tax systems very similar to the Condorcet winner, despite the differences in how poor individuals are weighted; the reason we get similarity is that poor individuals are already given high weight, so increasing their weight doesn't make much difference.

Obviously our experiments have limitations. Our policy space is restricted heavily, for feasibility reasons – we assume that households vote over time-invariant tax systems only, and they vote only once. To explore whether this commitment is crucial, we introduce a revote – at some point along the transition to the new steady state, we allow households to reconsider their tax system. We find that the Condorcet winner continues to win, and this result includes the new steady state as well. There are two reasons for this result. First, the Condorcet winner enjoys very strong support; in fact, it would survive under standard supermajority requirements. Second, the distribution of wealth does not change significantly along the transition, so agents are almost in the same situation as they started; as a result, the voting outcome is unchanged. Thus, we believe that commitment may not be critical, and given that abandoning it is very computationally burdensome we do not

explore this extension.¹⁰

We expect things may change if we expand the vote to permit time-varying tax systems. Dyrda and Pedroni (2014) find that the Ramsey optimal tax plan involves capital and labor income taxes that change over the transition to the new steady state, suggesting that more flexibility in the policy space could give us different answers.¹¹ Unfortunately, computational considerations prevent us from exploring this direction at this time (and for the foreseeable future).

1.1 Related Literature

Dolmas (2008) is the paper most closely related to this one, as he uses the same solution concepts; we discuss that paper explicitly in the appendix, rather than the main body of the paper, because the underlying economic environment is very different. Some papers that study the politics of multidimensional tax systems, such as Krusell, Quadrini, and Ríos-Rull (1996), set up their models to have a single decisive type whose identity is known. Others, such as Bassetto and Benhabib (2006) or Piguillem and Schneider (2013), have multi-dimensional voting but are able to prove the existence of a Condorcet winner; in effect, their models reduce to a single policy dimension and the median voter theorem applies. None of these papers have mobility, which is key for obtaining the analytical results.

We want to mention several papers that study political outcomes in the Aiyagari (1994) framework (and therefore have mobility). We have already mentioned Bachmann and Bai (2013), who examine how the government determines the supply of public goods. Corbae, D’Erasmus, and Kuruşçu (2009) study how a majority voting would select an income

¹⁰Multiple equilibria may be an issue without commitment, in the sense that the essential set for the current vote may be affected by beliefs about the essential set of future votes. There does not appear to be a computationally-feasible method for solving such a model.

¹¹Dyrda and Pedroni (2014) also permit the government to accumulate assets or issue debt, which has implications for the speed and direction of tax changes. The point we make here is merely that constant taxes are unlikely to be the preferred choice of any type.

tax/lump-sum transfer system and how that choice changes in response to rising idiosyncratic labor income risk. Aiyagari and Peled (1995) also study voting equilibria in a model of idiosyncratic risk, but they assume the economy instantly transits to the terminal steady state. All of these papers have single-dimensional policy spaces, so that they can apply the median voter theorem if preferences are single-peaked (which is difficult to verify in general).

There are a number of other papers that study political outcomes in the growth model. Rather than exhaustively list them here, we just note that our paper seems to be the only one that combines multi-dimensional voting with mobility and non-trivial decisive voter types¹².

2 Model

The model economy is populated by a continuum of *ex ante* identical, infinitely-lived households as in Aiyagari (1994). Each household receives an uninsurable shock ε to their wages in each period, which follows the process

$$\log(\varepsilon') = \rho \log(\varepsilon) + \sigma \zeta'$$

with $\zeta \sim N(0, 1)$ ¹³; these shocks are uncorrelated across individuals, implying that the distribution of ε can be assumed constant over time.¹⁴ Preferences are represented recursively as

$$v(k, \varepsilon) = \max_{c, h, k'} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \theta \log(1-h) + \beta E[v(k', \varepsilon') | \varepsilon] \right\}$$

¹²DeDonder (2000) also studies tax system determination using multi-valued solution concepts and finds, as we do, that the uncovered set is small (although he does not have a Condorcet winner in general). However, his model is static, so capital income taxation is not distortionary, and his parameters are not calibrated to match any cross-sectional facts. In contrast, our model is fully dynamic and we are explicit about the connection between our parameters and U.S. data.

¹³Here we follow the convention of recursive methods by denoting next period values with a prime.

¹⁴We invoke an appropriate law of large numbers, such as the one from Sun (2006).

subject to nonnegativity constraints and the budget constraint

$$(1 + \tau_c) c + k' \leq (1 + (1 - \tau_k)(r - \delta)) k + (1 - \tau_l) w \varepsilon h + T.$$

Here, c is consumption, k is the current holdings of assets, h is labor effort, r is the rental rate on capital, and w is the wage per efficiency unit of labor. The taxes are on consumption τ_c , labor income τ_l , and capital income τ_k , with lump-sum tax/transfer T .

The production sector consists of a single firm that operates the technology

$$Y = Z K^\alpha N^{1-\alpha}$$

where Z is total factor productivity, K is the aggregate capital stock, and N is aggregate labor input.¹⁵ The optimal factor demands implicitly satisfy the equations

$$\begin{aligned} r &= \alpha Z \left(\frac{K}{N} \right)^{\alpha-1} \\ w &= (1 - \alpha) Z \left(\frac{K}{N} \right)^\alpha. \end{aligned}$$

The government budget constraint is

$$G + T = \tau_c C + \tau_l w N + \tau_k (r - \delta) K,$$

where G is wasteful government spending.¹⁶

Markets clear if

$$\begin{aligned} K &= \int_k \int_\varepsilon k \Gamma(k, \varepsilon) \\ N &= \int_k \int_\varepsilon \varepsilon h(k, \varepsilon) \Gamma(k, \varepsilon) \\ C &= \int_k \int_\varepsilon c(k, \varepsilon) \Gamma(k, \varepsilon) \end{aligned}$$

and

$$C + K' - (1 - \delta) K + G = Y.$$

¹⁵ Z facilitates calibration, but otherwise merely serves to normalize units.

¹⁶We abstract from government debt partly for computational reasons and partly due to issues regarding the intertemporal solvency of the government.

2.1 How taxes matter

Households rank tax policies according to how their own wealth is affected. There are two aspects to the wealth effect of taxes – the direct effect and the indirect effect operating through changes in equilibrium prices. We discuss each in turn.

2.1.1 Direct effects

Setting aside general equilibrium effects, a household favors tax policies which place the lowest net tax burden on it, where the net tax burden is

$$T - \tau_c c - \tau_k (r - \delta) k - \tau_l w \varepsilon h.$$

This is achieved by a policy which places a low tax rate on the household's primary income source and a high rate on the other sources. For instance, households with high wealth work fewer hours and thus favor labor income taxes. Asset-poor households on the other hand derive most of their income from labor and so prefer to tax capital or consumption instead.

2.1.2 General equilibrium effects

In addition to this direct tax consideration, households also take into account how policies indirectly tax them through their effect on equilibrium prices. That is, they internalize the behavioral responses of other agents through substitution effects. The key conditions are the first-order conditions that describe trade-offs between (i) consumption and leisure today and (ii) consumption today and consumption tomorrow.

The intratemporal condition

$$\left(\frac{1 - \tau_{l,t}}{1 + \tau_{c,t}} \right) \frac{w_t}{c_t^\gamma} \leq \frac{\theta}{1 - h_t}$$

holds with equality when the non-negativity constraint on hours does not bind. Because the ratio $\frac{1 - \tau_{l,t}}{1 + \tau_{c,t}}$ can be the same for an infinite number (τ_c, τ_l) combinations, it can be tempting to view labor taxes and consumption taxes as equivalent (see Prescott 2004). This

equivalence holds in a representative agent environment, because there is no wealth effect and the representative household must supply positive labor hours. In our environment, in contrast, the two taxes are equivalent in terms of the substitution effect only for agents that work; for households that supply zero hours, local variations in taxes generate no substitution effect. Furthermore, since households differ in terms of their wealth, the wealth effects are not equivalent – it is easiest to see this point by noting that households who supply zero hours suffer no wealth loss from a labor tax but do lose from consumption taxation.

The intertemporal condition

$$\beta E_t \left\{ \left(\frac{1 + \tau_{c,t}}{1 + \tau_{c,t+1}} \right) \left(\frac{c_t}{c_{t+1}} \right)^\gamma [(1 - \tau_{k,t+1})(r_{t+1} - \delta) + 1] \right\} \geq 1$$

holds with equality if the household holds a positive amount of capital. Here, we see that consumption taxes can act as taxes on gross financial returns, if they fall over time; since we assume constant taxes, the intertemporal channel is shut down so that consumption taxes act only on initial wealth. Capital income taxes distort the consumption decision by reducing the return to saving and push consumption forward in time for all agents who are not borrowing constrained.

Ultimately, an individual household cares only about changes in r , w , and T which arise through changes in the aggregates C , K , and N . The total effect of individual behavior on these aggregates is unclear because it depends on the stationary distribution of wealth across households, and this distribution is an equilibrium outcome about which we can say little analytically (we know it exists and depends continuously on prices, but that's about all we can say). Next we calibrate and solve the model in order to quantify the cross-sectional distribution of wealth and substitution effects..

3 Calibration

We choose the parameters of our model to match some facts about the US economy. Specifically, we target a capital/output ratio of 12, a government/output ratio of 0.2, an investment/output ratio of 0.15, aggregate hours equal to 0.3, and capital's share of income equal to 0.36, along with a steady state output level of $Y = 1$. Combined with the US tax system of $(\tau_c, \tau_l, \tau_k) = (0.064, 0.234, 0.273)$ we obtain a transfer equal to 9.1 percent of GDP, quite close to the value estimated by Krusell and Ríos-Rull (1999). Table (1) presents the structural parameters that give rise to this calibration.

We set $\rho = 0.978$ and $\sigma = 0.0516$, consistent with the values from Flodén and Lindé (2001). Thus, wage shocks are very persistent and volatile, leading to significant welfare losses associated with incomplete insurance. Finally, we set $\gamma = 2$.¹⁷

Our model matches reasonably well the inequality and mobility observed in US data, but not perfectly. Figures (3a) and (3b) present the Lorenz curves for earnings and wealth in the US and in the model; while we do not get enough concentration in either we do relatively well. Tables (2) and (3) show mobility statistics from the model and the data over five-year horizons; again, we do surprisingly well, with clear deficiencies only in the extreme quintiles. In both cases, the model's main failures involve the extremes of the distribution. Under the benchmark of equal-weighted voting, these failures will not matter significantly since these extremes will have little mass; under wealth-weighted voting, however, they may play a more significant role.

We solve our model using standard methods – full details of the solution algorithm can be obtained upon request.

4 Majority Voting Solution Concepts

Our main experiment is to consider what tax systems a majority voting rule could select. Here, we call a tax system a vector $\tau = (\tau_c, \tau_l, \tau_k)$ along with a sequence of transfers $\{T_t\}_{t=0}^{\infty}$

¹⁷Setting $\gamma = 1$ did not materially affect our results.

that satisfy the government budget constraint period by period. We only allow tax systems that do not require lump-sum taxation in any period t , including the terminal steady state, so that $T_t \geq 0$ must hold. Additionally, we also do not permit negative tax rates.¹⁸

We define the payoff function

$$V(\tau; k, \varepsilon) = v(k, \varepsilon; \tau). \quad (1)$$

Note that we think of the payoff function as depending on taxes with (k, ε) as parameters, in contrast to the lifetime utility function v which reverses these variables. We can then define the preference relation $\mathcal{P}(k, \varepsilon)$ by

$$\tau_i \mathcal{P}(k, \varepsilon) \tau_j \Leftrightarrow V(\tau_i; k, \varepsilon) \geq V(\tau_j; k, \varepsilon).$$

Ties are unlikely in our model due to numerical approximation, so we simply ignore them. It is a straightforward use of the Theorem of the Maximum to obtain that V is continuous in τ , which is important for our solution concepts.

We suppose that politics is a series of pairwise competitions between tax systems. With more than one dimension in the policy space, the median voter theorem generally does not apply; furthermore, establishing the single-peakedness of preferences in our environment is difficult, so the median voter theorem may not even apply to a one-dimensional policy space. As a result, a Condorcet winner may not exist, and therefore the ordering of the competition may change the outcome (that is, agenda control has value). To avoid taking a stand on the details of the agenda (in particular, who gets to decide what the agenda is), we use an object from political economy called the **essential set** E as our solution concept (see Dutta and Laslier 1999); the essential set is the set of all policies played with positive probability in some mixed strategy Nash equilibrium.¹⁹ To compute this set, we

¹⁸Both assumptions are restrictive, as both transfers and labor income are very close to zero in the Condorcet winner. Labor subsidies are the endogenous voting outcome in Azzimonti, de Francisco, and Krusell (2008).

¹⁹The essential set is a generalization of the **bipartisan set** from Laffond, Laslier, and Le Breton (1993); the bipartisan set is the mixed strategy support when the mixed strategy equilibrium is unique, whereas

first define two other sets, the Pareto set P and the uncovered set U , which satisfy the inclusion ordering

$$E \subset U \subset P.$$

Because solving for E is computationally infeasible with a large policy space, we use P and U to eliminate many options before computing E . If a Condorcet winner exists, it will be the unique element of U and therefore also E .

The **Pareto set** P is defined as all tax systems that are not defeated unanimously by some other system. That is,

$$P = \left\{ \tau_i : \int_k \int_\varepsilon \mathbf{1}(\tau_j \mathcal{P}(k, \varepsilon) \tau_i) \Gamma_0(k, \varepsilon) < 1 \text{ for every } \tau_j \right\};$$

it is easy to see that the Pareto set cannot be empty and that no policy outside the Pareto set would be proposed (note that policies in the Pareto set still may never win under majority voting, and in pure redistribution games where total resources are fixed the Pareto set is the entire policy space). In practice, due to the very small measure of some types, we compute an "approximate Pareto set" where we set the criterion that the policy is not defeated by a vote of at least 0.9999. The approximate Pareto set in our model is contained in a relatively small section of the policy space.

Next, we say that τ_i **covers** τ_j if

$$\int_k \int_\varepsilon \mathbf{1}(\tau_i \mathcal{P}(k, \varepsilon) \tau_j) \Gamma_0(k, \varepsilon) > \frac{1}{2}$$

and

$$\int_k \int_\varepsilon \mathbf{1}(\tau_i \mathcal{P}(k, \varepsilon) \tau_q) \Gamma_0(k, \varepsilon) > \frac{1}{2}$$

for all τ_q such that

$$\int_k \int_\varepsilon \mathbf{1}(\tau_j \mathcal{P}(k, \varepsilon) \tau_q) \Gamma_0(k, \varepsilon) > \frac{1}{2}.$$

That is, τ_i covers τ_j if i beats j and i beats every policy q that j beats. The **uncovered set** U is then the set of tax systems that are not covered:

$$U = \{\tau_i : \nexists \tau_j \text{ such that } \tau_j \text{ covers } \tau_i\}.$$

the essential set is the union of the strategy support for all mixed strategy equilibria. See Dutta and Laslier (1999) for discussion.

Since some policies in the Pareto set may never win under majority voting, the uncovered set in general is a strict subset.

To compute the uncovered set, we construct the **adjacency matrix** M with (i, j) element

$$m_{i,j} = \begin{cases} 1 & \text{if } i \neq j \text{ and } \int_k \int_\varepsilon \mathbf{1}(\tau_i \mathcal{P}(k, \varepsilon) \tau_j) \Gamma_0(k, \varepsilon) > \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}.$$

We then compute the matrix

$$M^* = M^2 + M + I;$$

if $m_{i,j}^* = 0$ then i is covered by j . The uncovered set contains the policies that will survive under agenda setting with sophisticated voting, and as with the Pareto set is always nonempty; that is, any policy that could win under some particular ordering of pairwise competitions is contained in U .²⁰ A Condorcet winner is obviously uncovered, and is easily seen to be the unique element of U . If we continue by iteratively eliminating covered policies we obtain the **ultimately uncovered set** U^* (see Dutta 1988); if U^0 is the uncovered set given the Pareto set, then we can look for members of the uncovered set that are now covered given that some alternatives have been eliminated.²¹

To compute the essential set, we take all policies in the uncovered set. Let s_i denote the probability assigned to policy vector i , drawn from a set with $N < \infty$ possible policies. We construct the **dominance matrix**, with typical i, j element given by

$$d_{i,j} = \begin{cases} 1 & \text{if } i \neq j \text{ and } \int_k \int_\varepsilon \mathbf{1}(\tau_i \mathcal{P}(k, \varepsilon) \tau_j) \Gamma_0(k, \varepsilon) > \frac{1}{2} \\ 0 & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } \int_k \int_\varepsilon \mathbf{1}(\tau_i \mathcal{P}(k, \varepsilon) \tau_j) \Gamma_0(k, \varepsilon) < \frac{1}{2} \end{cases},$$

where $\Gamma_0(k, \varepsilon)$ is the initial distribution. We then solve the linear program

$$(s^*, e^*) = \underset{e, \{s_1, \dots, s_N\}}{\operatorname{argmax}} \{e\}$$

²⁰See McKelvey (1986) or Miller (2007) for discussion. The uncovered set is a nonempty subset of the **dominating set** (Smith 1973), in which all members of the set defeat all non-member policies in pairwise competition. In fact, the uncovered set is the smallest dominating set (Fishburn 1977).

²¹This process is related to the difference between a policy being locally covered versus being covered at a distance, as discussed in Miller (2007).

subject to

$$\begin{aligned}
\sum_{j=1}^N d_{i,j} s_j &\leq 0 \quad \forall i \in \{1, \dots, N\} \\
s_i - \sum_{j=1}^N d_{i,j} s_j - e &\geq 0 \quad \forall i \in \{1, \dots, N\} \\
\sum_{i=1}^N s_i &= 1 \\
s_i &\geq 0 \quad \forall i \in \{1, \dots, N\} \\
e &\geq 0.
\end{aligned}$$

The essential set then consists of the tax policies that are played with strictly positive probability:

$$E = \{\tau_i : s_i^* > 0\};$$

given the constraint set it is clear that E is nonempty. If there is a Condorcet winner, it will be the only member of the essential set.²²

5 Voting over Taxation

We employ the political equilibrium concepts from Section 4 to characterize the outcome of a "once-and-for-all" majority vote over future tax rates. We initialize the model in steady state under the tax policy $\tau^0 = (0.064, 0.234, 0.273)$; this vector is the US tax system as described by Carey and Rabesona (2002). A majority vote decides what taxes will be from period $t = 1$ onward. When evaluating the welfare consequences of each potential policy, households include the discounted utility experienced along the transition to the associated terminal steady state.

²²This linear program is discussed in Brandt and Fischer (2008). It is unclear from the literature whether the essential set is contained in the ultimately uncovered set in general.

5.0.3 The set of candidate tax policies

The space of candidate policies, \mathcal{P} , is constructed from a non-uniform grid consisting of 4,435 combinations of (τ_c, τ_l, τ_k) with τ_c between 0 and 190 percent, τ_l between 0 and 90 percent, and τ_k between 0 and 70 percent.²³ For each candidate, we solve for the terminal steady state and transition path starting from the initial distribution with fixed policy τ^0 . If there is a negative lump-sum transfer in the final steady state or in any period of the transition, then that policy is deemed infeasible and discarded from the policy set. The value functions obtained from solving each transition are used to construct the indirect utility function over tax policy described by equation (1).

We find that the Pareto set has relatively few members, and (more importantly) that the uncovered set (and therefore also the essential set) is a singleton. That is, there is a unique policy within the set which would garner a majority of votes in head-to-head competition with any other policy in the set. We denote this policy τ^C (for "Condorcet winner"), and it is $(\tau_c, \tau_l, \tau_k) = (0.16, 0.00, 0.19)$. Compared to τ^0 , the Condorcet winner places a much lower tax burden on labor and capital income taxes, and a significantly higher tax on consumption. The overall size of government is considerably smaller as well, as the transfer declines from 0.091 (9.10 percent of GDP) to 0.003 (0.27 percent of GDP).²⁴ Capital and output increase by 21.6 percent and 10 percent, respectively.²⁵

Note that our Condorcet winner is closest, in the Euclidean sense, to the tax system of South Korea – low labor taxation combined with significant capital income and consump-

²³Specifically, we combine a "coarse" grid and a "fine" grid. The coarse grid is 1600 policies constructed from combinations of 10 percent intervals in each tax dimension. We then identify the essential set from the coarse grid and place 1 percent intervals in the neighborhood of any policies in that set. The essential set from the coarse grid was a singleton (i.e., there was a Condorcet winner) so the fine grid is centered around that element.

²⁴We suspect that the transfer is actually zero in the terminal steady state, but it is prohibitively expensive to locate that point and check that it beats the Condorcet winner; obviously, it makes little difference.

²⁵We can show that the hyperplanes that divide the population evenly along each possible direction all intersect at the Condorcet winner; the graph, however, is three-dimensional and unreadable, so we omit it.

tion taxation. If we confine our policy space to only the observed country tax systems, South Korea indeed arises as a Condorcet winner in the restricted set. But a strong majority of voters would oppose shifting from the Condorcet winner to the South Korean system.

5.1 Preferences over tax policies

Generally, a household favors tax policies which place a lower tax burden on its primary source of income. For example, households with high wealth (who work fewer hours all else equal) broadly favor labor income taxes, while asset-poor households prefer consumption and capital income taxes. In addition to this direct tax consideration, households also take into account how policies indirectly tax them through their effect on equilibrium prices. In general equilibrium a higher τ_l reduces labor supply and the equilibrium interest rate. Because the tax incidence of higher labor income taxes still falls on households without labor income, rich households do not want to institute an extremely high tax rate on labor income. Poor households likewise internalize the effects of higher capital income taxes on the equilibrium wage. With these direct and indirect channels in mind, we now discuss each type of tax in turn and how households' preferences over each one differs across the wealth distribution.

5.1.1 Consumption taxes

Using (1), we hold fixed τ_l and τ_k , vary the consumption tax rate, and plot the welfare gain (in utils) from electing alternative policies with higher consumption taxes than the Condorcet winner. Specifically, for a household with initial wealth k and labor productivity ε , the welfare gain from enacting policy τ instead of τ' is $V(\tau; k, \varepsilon) - V(\tau'; k, \varepsilon)$. If the welfare gain is positive, then a household of type (k, ε) prefers τ to τ' .

In figures (4a) - (4c), the welfare gain is increasing in wealth, a reflection of the direct tax channel discussed above. Because the policy change is permanent, an increase in the consumption tax is equivalent to a positive tax on initial wealth. Households with high k

prefer no consumption tax, while poor households support very high levels of τ_c .

For households supplying positive hours, there is a second form of direct taxation in that a consumption tax is equivalent to a labor tax. Support for consumption taxes therefore naturally decreases in ε . Notice that the welfare gain curves converge as k increases and the curves of low ε types merge at lower wealth levels. As k increases the wealth effect on leisure pushes households to the corner of their hours decision. Households with more labor productivity hit the lower bound on hours at a higher level of wealth because leisure is more expensive for them.

For households not supplying labor, the consumption tax/labor income tax equivalence matters only through the risk that the household will supply hours in the future. Typically in order for that to happen, the household must draw a sufficiently long string of below average productivity, the likelihood of which depends upon the persistence of the ε process. Consider a household with $\varepsilon = \varepsilon_{\min}$ and enough wealth so that the household does not work. First, the household will consume more than its after-tax capital income so it will have less wealth next period. Second, because productivity shocks are serially correlated, the household will very likely draw another low ε tomorrow, causing its wealth to decline further towards a level where the household would work. Finally, because higher productivity households exit the labor market at higher wealth levels, as the household reduces its wealth the probability that it will supply positive hours due to a favorable ε shock rises.

5.1.2 Capital income taxes

Figures (5a) - (5c) plot the welfare gains associated with varying the capital income tax. Support for increasing capital taxes falls in k and in ε . Declining support in k is obvious, but not necessarily so for ε . Why do only a minority of households with zero wealth vote against raising capital income taxes from the Condorcet winner level? There are two reasons. First, these households get most of their income from labor. Thus, the wage matters a lot to them, and high capital income taxes reduce the capital-labor ratio, driving down the wage; furthermore, wage increases benefit high ε types more than low ε types.

Second, there is a lot of mobility in the model, meaning that households transition through the wealth distribution quickly relative to their discount factor. Simply put, even zero wealth households expect to have a lot more wealth in the "near" future, and so they share much of the same distaste for capital income taxes as their currently rich counterparts.

5.1.3 Labor income taxes

The model predicts zero labor income taxes.²⁶ Preferences for increasing labor income taxes are monotonic decreasing in ε . Households that are more productive dislike labor taxes more, especially at low wealth levels where labor income is most important for consumption. However, labor income tax preferences over k are not monotonic. Instead each ε -curve is U-shaped over k and divided into three regions (see Figures (6a)-(6c)).

In the first region, households have low wealth and support low labor income taxes. As households gain wealth, they reduce their dislike of labor income taxes with low ε households being the first to move into the second region and vote for increasing the tax. The downward slope is once again due to the wealth effect on hours. Each ε type curve bottoms out near the wealth level where it supplies 0 hours. Notice the same merging of curves from the bottom first, as the corner is hit at higher wealth levels for higher ε due to the substitution effect.

The reason for that households eventually switch back to preferring lower labor income taxes is the effect of these taxes on the interest rate. Sufficiently-rich households do not work and do not expect to work for a long time. The income and consumption of these types depend entirely upon the return to capital. Labor income taxes increase the capital-labor ratio and reduce r , so rich households oppose them. The fraction of households in the second and third regions is small, however, so the outcomes are primarily determined by the first region.

²⁶Although we do not permit labor subsidies, it is likely that a policy with negative labor income taxes would defeat the Condorcet winner here. Azzimonti, de Francisco, and Krusell (2009) also find majority support for labor subsidies, but in a environment without capital and idiosyncratic risk.

5.1.4 Most preferred tax policy by household type

By plotting (1), we can see the most preferred tax combination across wealth by productivity level. Figures (7a) - (7c) do this for the lowest productivity, the median productivity, and the highest productivity households. For the lowest- ε households, the ideal tax combination is high consumption taxes, moderate capital income taxes, and zero labor income taxes. As wealth increases the preferred combination switches to zero consumption taxes, high labor income taxes, and slightly lower capital income taxes. The switch occurs around a wealth level at which the household earns a large fraction of its income from capital.

The picture is generally the same for all ε types. At low levels of k , consumption taxes are preferred to labor taxes with capital income somewhere in between. Then at a higher level of wealth the ordering switches so labor income taxes become preferred to consumption taxes. The differences across types are mainly the degree to which they wish to tax consumption when at low k levels. The least productive wealth-poor want very high consumption taxes while their most productive counterparts want moderate levels. The k level at which the preferences switch increases with ε since the substitution effect on hours is stronger, keeping high ε types working more, and thus more exposed to labor income taxation.

The figures also show the cumulative distribution of wealth. By taking the distribution of voters along with their preferences, we can understand the voting outcome. Labor income taxes are opposed by the majority voting system not because no one favors them, but because essentially no households reach the wealth levels where they become preferred. We show in a later section that wealth-weighted voting changes equilibrium policy to taxing labor.

The Condorcet winner places no direct tax on labor and a mixture of taxes on consumption and capital income. It is natural to wonder why the Condorcet winner features both of these taxes rather than a single tax on consumption. The decisive voter balances the regressive nature of the consumption tax against the distortion on saving (and therefore on the wage) of the capital income tax. The consumption tax is regressive because

the marginal propensity to consume out of income declines in wealth, meaning that the marginal cost of consumption taxation falls more heavily on low-wealth households; naturally, then, they oppose imposing large consumption tax rates. In contrast, the capital income tax is progressive, but comes with a substantial cost in terms of aggregate wages – as seen in Figure 1 in Aiyagari and McGrattan (1998), the economy operates in general on the highly-elastic portion of the capital supply curve, so that even small changes in after-tax returns will lead to large changes in aggregate capital and therefore wages. Thus, voters are reluctant to impose high capital taxes; the result is they settle for a compromise that mixes the two types. We show later, in a section on probabilistic voting, that the voters are actually roughly equating the marginal net gains from the two taxes.

We can see further how unpopular labor taxes are by considering what would happen if we forced capital and labor to be taxed at a uniform rate (an income tax). Here, the Condorcet winner chooses a consumption tax rate of 20 percent and zero income taxation. This result is consistent with other papers that show how desirable consumption taxation is, compared to income taxation; our results show that it is missing the key point that income taxation per se is not unpopular, only labor income taxation. If we remove consumption taxation as a policy instrument, forcing households to finance government expenditures with taxes on labor and capital income, they will tax labor but only because to finance with only capital income taxes requires an extremely high tax rate which reduces too much.

5.1.5 Support for the Condorcet winner

We now consider how strong the support is for the winning policy. If one considers an alternative policy which deviates from the Condorcet winner along a single dimension only, the share of votes going to that alternative declines in the size of the deviation. Figure (8a) plots the fraction of households voting against the Condorcet winner for alternatives with higher consumption tax rates. Support for the alternative decreases sharply as τ_c increases. The Condorcet winner has a comfortable majority even against $\tau_c = 0.3$, with over 61 percent of the vote. Figure (8b) illustrates that even small deviations from zero

labor income taxation are met with little support. Figure (8c) shows the same monotonic decline in support as policy moves away from the Condorcet winner in the capital income tax direction. The vote share is especially sensitive in this dimension.

5.1.6 Transition Path

Figures (9) - (11) plot the transition paths of economy aggregates resulting from a change to the fine grid Condorcet winner. There is a decline in most aggregates in the very early periods of transition. The capital stock declines for 5 periods before bottoming out at 2 percent below initial steady state. Over time it rises reaching a level 0.5 percent higher at the end of the transition. Effective labor input, hours, consumption, and output all fall sharply initially and never rise back to their initial levels. Despite considerable initial wealth and income inequality, majority voting in the model does not lead to much redistribution in the long run. The transfer falls from just over 9 percent of GDP in the initial steady state to 6 percent in the first period of transition before declining further to nearly 0 and remaining there for the rest of the transition.²⁷

5.1.7 Inequality and Revoting

The Condorcet winner gives rise to a long run wealth distribution with more inequality than the initial distribution. This change can be seen from the Lorenz curves of wealth plotted in Figure (12). The initial distribution has Gini coefficient of 0.60, while the Condorcet distribution is only slightly higher at 0.62. We interpret these small changes as evidence that our assumption of a once-and-for-all vote is unlikely to be restrictive; the agents will be in nearly the same distribution in each time period and therefore will not generate substantial support for alternative policies (particularly since the Condorcet winner enjoys very strong support). To check this intuition, we repeat the vote using the wealth distribution at different points in the transition, namely periods $t = 30$, $t = 60$,

²⁷Experimenting with capital adjustment costs led to no meaningful changes, only a correspondingly slower transition. The Condorcet winner remained the same.

and the terminal steady state. In each case, the Condorcet winner continued to exist and enjoy strong support.²⁸

Because there are mixed effects on inequality from changing the various tax rates, we once again isolate tax changes along a single dimension and compare long-run wealth distributions. Figure (13) plots the cdf's of the steady state wealth distributions for alternative policies with much higher consumption taxes. As τ_c rises, long run inequality is unambiguously reduced, which is the wealth tax equivalence of consumption taxation appearing again since higher consumption taxes redistribute initial wealth. Figure (14) plots cdf's for changes in capital income taxes. Just like consumption taxes, capital income taxes also unambiguously decrease long run wealth inequality. Labor income taxes have a very small effect on the lower end of the wealth distribution, increasing inequality just slightly. At the middle and upper ends, however, labor income taxes have much stronger inequality-reducing effects. The cdf's are plotted in Figure (15).

The reasons for the small effect on inequality can be traced back to the large amount of idiosyncratic risk faced by agents. If risk is small, changes in factor prices play a large role at determining inequality – in the limit where risk is zero and households are frozen in their place in the wealth distribution, small changes in the interest rate lead to large changes in absolute wealth positions. But with idiosyncratic risk this effect gets swamped by precautionary motives, and changes in flat taxes simply do not materially affect this dimension.²⁹

²⁸This procedure assumes that agents do not anticipate an opportunity to revote. It is therefore not necessarily equivalent to assuming the vote occurs each period – computing an equilibrium without commitment is complicated by the potential for multiple solutions, so we leave it for future work. One-dimensional models of voting without commitment in similar environments are studied in Bachmann and Bai (2013) and Corbae, D'Erasmo, and Kuruşçu (2009), where a "Krusell-Smith"-style approximation is used to find approximate equilibria.

²⁹Carroll and Young (2016a) discuss the connection between inequality and factor prices in a related environment, and show that idiosyncratic risk dominates any other factors once it is sufficiently important.

5.1.8 The Role of Mobility and Patience

Our results depend critically on two aspects of our economy – households are mobile (they move around the (k, ε) space) and they are very patient. We illuminate the role of mobility by studying a version of our economy in which ε is fixed at the current value; as noted by Krusell and Ríos-Rull (1999) and Chatterjee (1994), the resulting economy has no mobility.³⁰ What we find is that the level of transfers is substantially higher, along with the levels of capital and consumption taxation, but labor income taxes remain zero. Mobility matters because agents expect to be "average" in the future; combined with a high β they become reluctant to impose high taxes since they are committed to them for all time. Tables (2) and (3) show that our model actually overstates mobility for certain groups (based on mobility measures from Budría Rodríguez *et al.* 2002), namely the very wealthy and very poor; with less mobility we would find that support for higher transfers would increase, although given the strong majority enjoyed by the Condorcet winner it seems unlikely that it would rise enough to rationalize the US level.³¹

To examine the role of patience (specifically, the forward-looking nature of voting), we consider a case where households evaluate the transition for the purposes of voting using a discount factor $\beta^* < \beta$; we call this "myopic voting". That is, households make economic decisions using

$$v_t(k, \varepsilon; \tau) = \frac{c_t(k, \varepsilon)^{1-\gamma}}{1-\gamma} + \theta \log(1 - h_t(k, \varepsilon)) + \beta E[v_{t+1}(k_{t+1}(k, \varepsilon), \varepsilon'; \tau) | \varepsilon]$$

but vote using

$$\tilde{v}_t(k, \varepsilon; \tau) = \frac{c_t(k, \varepsilon)^{1-\gamma}}{1-\gamma} + \theta \log(1 - h_t(k, \varepsilon)) + \beta^* E[\tilde{v}_{t+1}(k_{t+1}(k, \varepsilon), \varepsilon'; \tau) | \varepsilon].$$

We find that the transfer again rises substantially, and both capital and consumption taxation increase. The reason is that now the future is discounted more, meaning that

³⁰Specifically, each household's share of wealth remains equal over time: $\frac{k_t}{K_t} = \frac{k_0}{K_0}$. Note however that the decisive voter is still endogenously determined and could change over the transition, so commitment to policies is still an issue.

³¹Carroll and Young (2016b) investigate the reasons why the model misses in the extreme quintiles.

the currently "poor" are less affected by their transition to being average in the future. However, labor income taxes remain at zero. Depending on the degree of myopia in voting, we can generate very large transfers as a Condorcet winner. A complete description of these experiments is available upon request.

5.2 Wealth-Weighted Voting

In the experiment above, voting is proportional: the vote share cast by households in state (k, ε) is equal to the population share at (k, ε) . Bachmann and Bai (2013) document that political participation (in the form of campaign contributions) varies substantially across the wealth distribution, with participation rising in wealth. They propose a weighting function, k^χ , which multiplies the population weights, where χ captures the degree of "wealth bias" in the political process. There are multiple interpretations of this bias – 'One-Dollar-of-Contribution-One-Vote' implies that one dollar of contribution buys one de facto vote, while 'One-Contributor-One-Vote' means that each contributor (independent of dollar amounts) gets one de facto vote. Bachmann and Bai (2013) use data from the Cooperative Congressional Election Survey and the American National Election Studies to discipline their choices of χ ; when $\chi = 0.55$ they find that the model matches reasonably well the political bias generated by contributions and matches the low correlation between government spending and GDP.³²

We apply wealth-weighting to our model and find that it moves equilibrium tax policy away from capital income and consumption taxation and toward labor taxation. When $\chi = 0.3$, the winning policy is $(\tau_c, \tau_l, \tau_k) = (0.0, 0.20, 0.10)$; at $\chi = 0.5$ we obtain $(\tau_c, \tau_l, \tau_k) = (0.0, 0.30, 0.0)$. To get a better sense of how much wealth bias is implied by these χ values, Figure (16) plots the cdf's of votes cast as a function of wealth under the two wealth-weighting rules and under proportional voting ($\chi = 0$). When votes are counted

³²Bai and Lagunoff (2013) show formally the positive link between wealth inequality and political inequality for this weighting function. Note that the interpretation is a bit strained, since agents in the model don't actually sacrifice any resources in order to gain voting power.

proportionally, the 50th percentile of votes occurs at the median wealth level, $k = 6.6$. In contrast, the 50th percentile of votes is at the 69th and 75th percentiles of the $\chi = 0.3$ and $\chi = 0.5$ wealth distributions, respectively.

It is easy to understand these changes – the rich like labor taxes, as shown above, and when they are sufficiently important politically they impose this preference. The point we make here is that (i) it is easy and feasible to introduce aspects of voter participation and (ii) voter participation may play an important role in cross-country comparisons of tax systems.

5.3 Probabilistic Voting

Deterministic voting in multi-dimensional environments may need additional assumptions about the structure of the political process for equilibrium selection. This was not a significant concern in our analysis; however because of this concern, many researchers have adopted an alternative structure for majority voting – probabilistic voting. Under probabilistic voting, a household's return from any particular tax policy is a function of two elements: the deterministic economic component $V(\tau; k, \varepsilon)$ and a "non-economic," policy-specific shock. As a result, given a distribution for those shocks the expected number of votes for each option is continuous. Supposing that the objective of the political candidates is to maximize the probability of winning, this continuity guarantees the existence of an equilibrium.

While the precise function for voter utility is arbitrary, in practice the distribution of shocks is logistic and enters utility either additively or multiplicatively; for these cases expected vote totals can be calculated analytically. Banks and Duggan (2005) show that for either assumption about the way shocks enter the payoff function, there is an equivalence between solving the probabilistic voting game and maximizing a social welfare function. When shocks enter additively, the social welfare function is the sum of households' utility (the utilitarian criterion). When the shocks enter multiplicatively, then the social welfare function is the sum of the logarithm of utility, leading to a social planner that displays

constant relative inequality aversion (see Atkinson 1970), meaning that low-utility types get disproportionate weight in the objective function.³³

To implement probabilistic voting, we assume that agents receive an additive shock η_i to the utility of τ_i when voting; that is, the payoff from i is $V(\tau_i; k, \varepsilon) + \eta_i$. Household (k, ε) would vote for policy i over policy j if

$$V(\tau_i; k, \varepsilon) + \eta_i > V(\tau_j; k, \varepsilon) + \eta_j,$$

which can be rearranged to obtain

$$V(\tau_i; k, \varepsilon) - V(\tau_j; k, \varepsilon) > \eta_j - \eta_i.$$

Supposing η has a logistic distribution independent of (k, ε) , i competing against j will receive an expected vote share equal to

$$2 \int_k \int_\varepsilon \frac{\exp(V(\tau_i; k, \varepsilon))}{\exp(V(\tau_i; k, \varepsilon)) + \exp(V(\tau_j; k, \varepsilon))} \Gamma_0(k, \varepsilon) - 1.$$

Alternatively, we can assume that the shock is multiplicative, so that the payoff is $\exp(\eta_i) V(\tau_i; k, \varepsilon)$, which leads to an expected vote share of

$$2 \int_k \int_\varepsilon \frac{V(\tau_i; k, \varepsilon)}{V(\tau_i; k, \varepsilon) + V(\tau_j; k, \varepsilon)} \Gamma_0(k, \varepsilon) - 1.$$

The preferred policy maximizes the expected number of votes, evaluated at a symmetric equilibrium $\tau_i = \tau_j$.

We find that the additive case picks $(\tau_c, \tau_l, \tau_k) = (0.18, 0.00, 0.14)$ and the multiplicative case $(\tau_c, \tau_l, \tau_k) = (0.19, 0.00, 0.12)$. Both outcomes are very similar to the Condorcet winner, showing that the political process delivers a good outcome (at least in terms of average welfare). This result allows us to give a welfare interpretation to why voters choose

³³In general probabilistic voting and deterministic majority voting deliver different outcomes; see as a comparison case Hassler et al. (2003) vs. Hassler et al. (2005). Dolmas (2014) points out a potential pitfall from using probabilistic voting: in some environments, the outcome depends critically on whether the shocks are assumed to enter additively or multiplicatively. Since the shocks are non-economic, no data can ever be directly brought to bear on them.

both consumption and capital income taxes. Lipsey and Lancaster (1956) note that, in general, second-best outcomes use multiple instruments a small amount rather than one instrument a lot; the intuition is that deadweight losses from taxation are typically convex in the tax rate. In this context, voters like consumption taxes because they act as a tax on initial wealth, but unfortunately it is a regressive tax – the poor suffer disproportionately more because their marginal propensity to consume is high. Voters like capital income taxes because they tax the wealthy more heavily than the poor, but at a cost of permanently lower aggregate wages. Therefore, the voters compromise and use a little bit of both; using the welfare maximization interpretation, the mix of taxes equalizes the marginal gains net of costs.

We can also explain why the additive and multiplicative results are very similar to the Condorcet winner, despite the differences in the social welfare function that they maximize. Relative to the utilitarian case, the inequality-averse welfare function upweights the utility of poor agents at the expense of rich ones. But the political outcomes already weigh these agents significantly, since they are numerous enough to be decisive under majority voting.

6 Conclusion

In our study of political mechanisms in the Aiyagari (1994) model, we found that majority voting was quite powerful at restricting outcomes – the Condorcet winner not only existed but enjoyed very strong support against alternatives. Compared to the data, however, the equilibrium tax system was far afield – although closest to the South Korean system of low labor income taxes and modest capital income and consumption taxes, a strong majority in the model would oppose shifting to South Korean taxes. Transfers are close to zero for nearly the entire transition and in the new steady state; in only one period is there a significant positive transfer. We interpret these results positively – if we want to understand the political choices made by countries with respect to their tax systems, it seems more fruitful to look at how the countries are different in terms of their endowments,

markets, and preferences rather than the intricate details of their political processes.

There are obviously many directions along which the calibration of the model could be improved in terms of earnings and wealth concentration, in particular by being more careful about the idiosyncratic wage process. For example, Guvenen et al. (2016) find evidence that log labor income changes are highly non-Gaussian. Guvenen et al. (2016) model their process as a mixture of AR(1) processes, with mixture weights that depend on lagged income and heterogeneous variances; it would be a challenge to introduce the full scope of their process into our model (notwithstanding the problem that they estimate earnings rather than wages), but some features could perhaps be added that increase the amount of heterogeneity and introduce wealth concentration. As we have noted already, mobility is also a concern given its importance at determining the level of transfers; Carroll and Young (2016) explore the relationship between the basic Bewley-Aiyagari model's predictions about wealth mobility and those found in PSID data. Human capital accumulation may also be important, although it seems likely that it would simply make labor taxation even more unattractive.

The machinery developed here can be used in a number of ways. Of particular interest to us would be the politics of progressive taxation, as the flexibility of a progressive tax system could allow the decisive block of poor and middle class voters to tax the rich without also having to tax themselves. The diversity of progressive tax systems in the data is documented by Holter, Krueger, and Stepanchuk (2014); we are exploring whether the model still predicts only a small number of policies could survive and whether transfers are politically feasible. We could also study the importance of tax avoidance (deliberately shifting out of taxable activities like market consumption into untaxed home consumption) and tax evasion (choosing not to pay taxes that are owed) for the determination of equilibrium tax systems. There is a large literature on the implementation costs of different taxes (that dates back at least to Yitzhaki 1979), and surely these costs play a role in the voting choices of agents. Our framework can easily accommodate such costs.

7 Appendix

In this Appendix we discuss how our results relate to Dolmas (2008). In that paper, majority voting was found to have limited predictive power – the uncovered set and the essential set both contained a large number of policies that were ”spread out” across the policy space, so that the old adage that ”in politics anything goes” held. We illustrate that this limited power was a result of one crucial assumption, namely that the number of distinct types was small.

The model of Dolmas (2008) is substantially different than ours, so we quickly describe the relevant ingredients. The economic model is taken from Rebelo (1991), and features endogenous growth and no transitional dynamics. Households are endowed with an initial wealth k_0 and productivity e_0 ; these values will remain fixed over time (in the case of wealth, fixed as a share the aggregate to be more precise, but since the model is always on a balanced growth path the distinction is irrelevant). Dolmas (2008) supposes there are two values for productivity and two for wealth, leading to four types, and calibrates their weights to roughly match some distributional facts from the US. The result is that no group is a majority, but each has substantial measure; in fact, one can assemble a majority in his model in eight different ways, most but not all of which include the ’low e , low k ’ type that has measure just smaller than 0.5. Each of these types has very different preferred tax rates, which depend on where their primary source of income arises and how wealthy they are relative to the mean; however, most types are poorer than average so they desire positive transfers, and because there is no mobility they are not reluctant to impose high taxes to get them. Figure (17) presents the uncovered set from Dolmas (2008) for his benchmark model with four types; note that the range of taxes is very large in each dimension.³⁴ As a comparison we also present in Figure (17) the uncovered set when the number of types is extended to nine, using a similar calibration procedure that picks points from the Lorenz curves for earnings and wealth in the US; now the uncovered set

³⁴A slight difference in the parameters accounts for the change in the location of the Pareto set relative to the graph in Dolmas (2008).

has shrunk to a small number of points (6) and is confined to a region of zero labor income taxes, high capital and consumption taxes, and very large transfers, and the essential set contains only 5 points. The Pareto set, in comparison, remains large; the specifics of these calculations are available upon request. This experiment shows that again our zero labor income tax result is very robust and the importance of mobility for the level of transfers. Since the model is very different from ours we do not explicitly compare the two.

References

- [1] Aiyagari, S.R., (1994). "Uninsured idiosyncratic risk and aggregate saving." *Quarterly Journal of Economics* 109, 659-684.
- [2] Aiyagari, S.R., and E.R. McGratten (1998) "The optimum quantity of debt." *Journal of Monetary Economics* 42, 447-469.
- [3] Aiyagari, S.R., and D. Peled (1995). "Social insurance and taxation under sequential majority voting and utilitarian regimes." *Journal of Economic Dynamics and Control* 19, 1511-1528.
- [4] Atkeson, A., V.V. Chari, and P.J. Kehoe (1999). "Taxing capital income: a bad idea." Federal Reserve Bank of Minneapolis *Quarterly Review* 23, 3-17.
- [5] Atkinson, A.B. (1970). "On the measurement of inequality." *Journal of Economic Theory* 2, 244-263.
- [6] Azzimonti, M., E. de Francisco, and P. Krusell (2008). "Production subsidies and redistribution." *Journal of Economic Theory* 142, 73-99.
- [7] Bachmann, R., and J. Bai (2013). "Politico-economic inequality and the comovement of government purchases." *Review of Economic Dynamics* 16, 565-580.
- [8] Bai, J., and R. Lagunoff (2013). "Revealed political power." *International Economic Review* 54, 1085-1115.

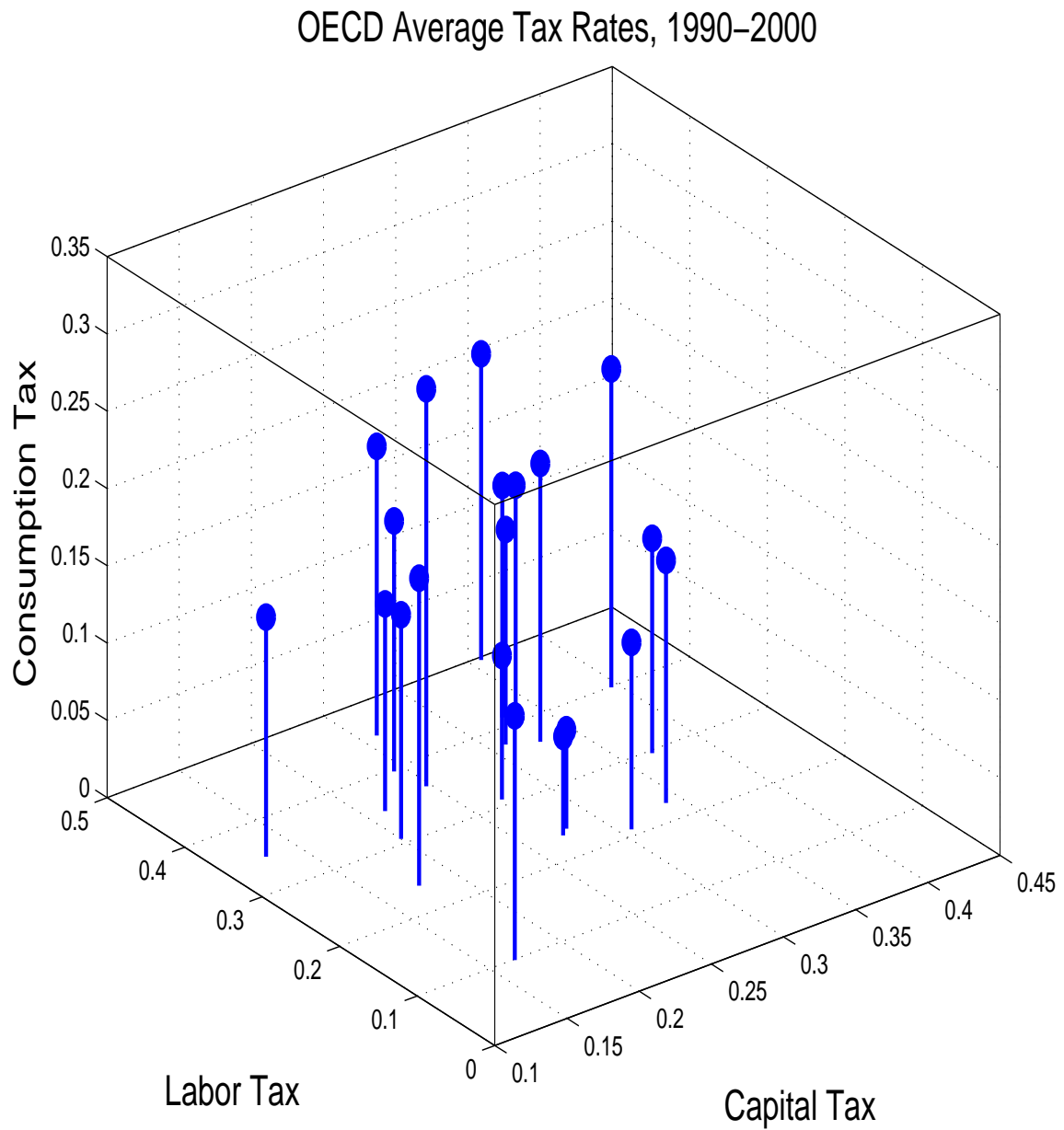
- [9] Banks, J.S., and J. Duggan (2005). "Probabilistic voting in the spatial model of elections: the theory of office-motivated candidates." In Austen-Smith, D., and J. Duggan (eds.), *Social Choice and Strategic Decisions*, Studies in Social Choice and Welfare, 15-56.
- [10] Bassetto, M., and J. Benhabib (2006). "Redistribution, taxes, and the median voter." *Review of Economic Dynamics* 9, 211-223.
- [11] Brandt, F., and F. Fischer (2008). "Computing the minimal covering set." *Mathematical Social Sciences* 56, 254-268.
- [12] Budría Rodríguez, S., J. Díaz-Giménez, V. Quadrini, and J.-V. Ríos-Rull (2002). "Updated facts on the U.S. distributions of earnings, income, and wealth." Federal Reserve Bank of Minneapolis *Quarterly Review* 26, 2-35.
- [13] Carey, D., and J. Rabesona (2002). "Tax ratios on labour and capital income and on consumption." *OECD Economic Studies* 35, 129-174.
- [14] Carroll, D.R., and E.R. Young (2011). "The long run effects of changes in tax progressivity." *Journal of Economic Dynamics and Control* 35, 1451-1473.
- [15] Carroll, D.R., and E.R. Young (2016a). "Neoclassical inequality." Manuscript.
- [16] Carroll, D.R., and E.R. Young (2016b). "Mobility in the Bewley model." Manuscript.
- [17] Caselli, F., and J. Ventura (2000). "A representative consumer theory of distribution." *American Economic Review* 90, 909-926.
- [18] Chari, V.V., J.P. Nicolini, and P. Teles (2016). "More on the optimal taxation of capital." Manuscript.
- [19] Chatterjee, S., (1994). "Transitional dynamics and the distribution of wealth in a neoclassical growth model." *Journal of Public Economics* 54, 97-119.
- [20] Chien, Y., and J. Lee (2006). "Why tax capital?" Manuscript.

- [21] Conesa, J.C., S. Kitao, and D. Krueger (2008). "Taxing capital? Not a bad idea after all!" *American Economic Review* 99, 25-48.
- [22] Corbae, P.D., P. D'Erasmo, and B. Kuruşçu (2009). "Politico-economic consequences of rising wage inequality." *Journal of Monetary Economics* 56, 43-61.
- [23] Correia, I. (2010). "Consumption taxes and redistribution." *American Economic Review* 100, 1673-1694.
- [24] Dávila, J., J.H. Hong, P. Krusell, and J.-V. Ríos-Rull (2012). "Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks." *Econometrica* 80, 2431-2467.
- [25] DeDonder, P., (2000). "Majority voting solution concepts and redistributive taxation." *Social Choice and Welfare* 17, 601-627.
- [26] Dolmas, J., (2008). "What do majority-voting politics say about redistributive taxation of consumption and factor income? Not much." Federal Reserve Bank of Dallas Working Paper 0814.
- [27] Dolmas, J., (2014). "Almost orthogonal outcomes under probabilistic voting: a cautionary example." Manuscript.
- [28] Dutta, B., (1988). "Covering sets and a new Condorcet choice correspondence." *Journal of Economic Theory* 44, 63-80.
- [29] Dutta, B., and J.-F. Laslier (1999). "Comparison functions and choice correspondences." *Social Choice and Welfare* 16, 513-532.
- [30] Dyrda, S., and M.Z. Pedroni (2014). "Optimal fiscal policy in a model with uninsurable idiosyncratic shocks." Manuscript, University of Minnesota.
- [31] Fishburn, P.C. (1977). "Condorcet social choice functions." *SIAM Journal of Applied Mathematics* 33, 469-489.

- [32] Flodén, M., and J. Lindé (2001). "Idiosyncratic risk in the U.S. and Sweden: is there a role for government insurance?" *Review of Economic Dynamics* 4, 406-437.
- [33] Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003). "Optimal indirect and capital taxation." *Review of Economic Studies* 70, 569-587.
- [34] Guvenen, F., F. Karahan, S. Ozdan, and J. Song (2016). "What do data on millions of U.S. workers reveal about life-cycle earnings risk?" Manuscript.
- [35] Hassler, J., P. Krusell, K. Storesletten, and F. Zilibotti (2005). "The dynamics of government." *Journal of Monetary Economics* 52, 1331-1358.
- [36] Hassler, J., J.V. Rodríguez Mora, K. Storesletten, and F. Zilibotti (2003). "The survival of the welfare state." *American Economic Review* 93, 87-112.
- [37] Holter, H.A., D. Krueger, and S. Stepanchuk (2014). "How does tax progressivity and household heterogeneity affect Laffer curves?" CFS Working Paper No. 490.
- [38] Krusell, P., and J.-V. Ríos-Rull (1999). "On the size of U.S. government: political economy in the neoclassical growth model." *American Economic Review* 89, 1156-1181.
- [39] Krusell, P., V. Quadrini, and J.-V. Ríos-Rull (1996). "Are consumption taxes really better than income taxes?" *Journal of Monetary Economics* 37, 475-503.
- [40] Laffond, G., J.-F. Laslier, and M. Le Breton (1993). "The bipartisan set of a tournament game." *Games and Economic Behavior* 5, 182-201.
- [41] Lipsey, R.G., and K. Lancaster (1956). "The general theory of second best." *Review of Economic Studies* 24, 11-32.
- [42] McKelvey, R., (1976). "Intransitivities in multidimensional voting models and some implications for agenda control." *Journal of Economic Theory* 12(3), 472-482.

- [43] McKelvey, R., (1986). "Covering, dominance, and institution-free properties of social choice." *American Journal of Political Science* 30, 283-314.
- [44] Mendoza, E.G., A. Razin, and L.L. Tesar (1994). "Effective tax rates in macroeconomics: cross-country estimates of tax rates on factor incomes and consumption." *Journal of Monetary Economics* 34, 297-323.
- [45] Miller, N.R., (2007). "In search of the uncovered set." *Political Analysis* 15, 21-45.
- [46] Piguillem, F., and A. Schneider (2013). "Heterogeneous labor skills, the median voter, and labor taxes." *Review of Economic Dynamics* 16, 332-349.
- [47] Plott, C.R., (1967). "A notion of equilibrium and its possibility under majority rule." *American Economic Review* 57, 787-806.
- [48] Prescott, E.C., (2004). "Why do Americans work so much more than Europeans? " *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2-13.
- [49] Rebelo, S., (1991). "Long-run policy analysis and long-run growth " *Journal of Political Economy* 99, 500-521.
- [50] Smith, J.H. (1973). "Aggregation of preferences with variable electorates." *Econometrica* 41, 1027-1041.
- [51] Straub, L., and I. Werning (2014). "Positive long-run capital taxation: Chamley-Judd revisited." *Working Paper No. w20441. National Bureau of Economic Research*.
- [52] Sun, Y. (2006). "The exact law of large numbers via Fubini extension and characterization of insurable risks." *Journal of Economic Theory* 126, 31-69.
- [53] Trabandt, M., and H. Uhlig (2011). "The Laffer curve revisited." *Journal of Monetary Economics* 58, 305-327.
- [54] Yitzhaki, S. (1979). "A note on optimal taxation and administrative costs." *American Economic Review* 69, 475-480.

Figure 1: Tax systems across the world

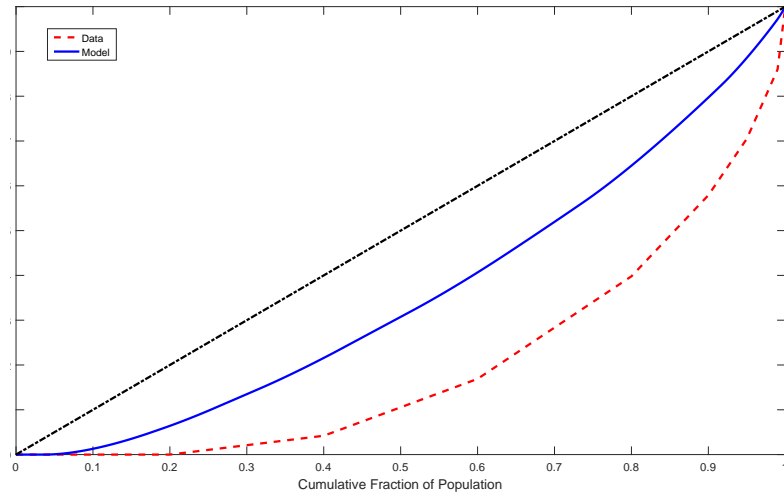


Source: Carey and Rabesona (2002)

Figure 2: *

Figure 3: Lorenz curves

(a) Earnings



(b) Wealth

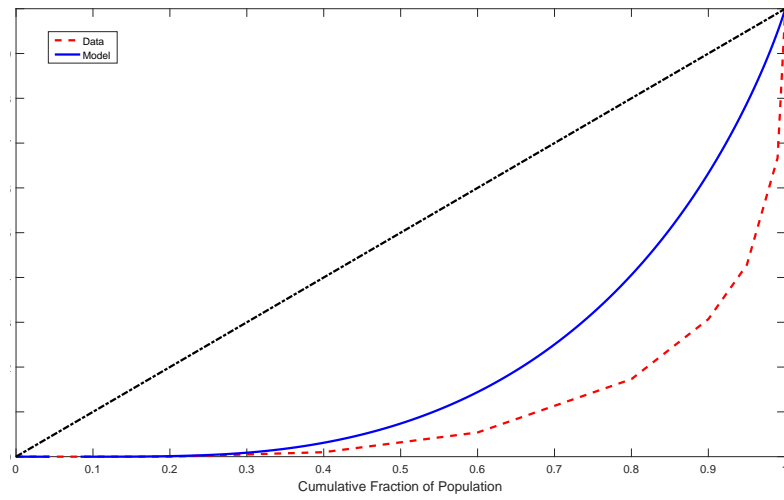
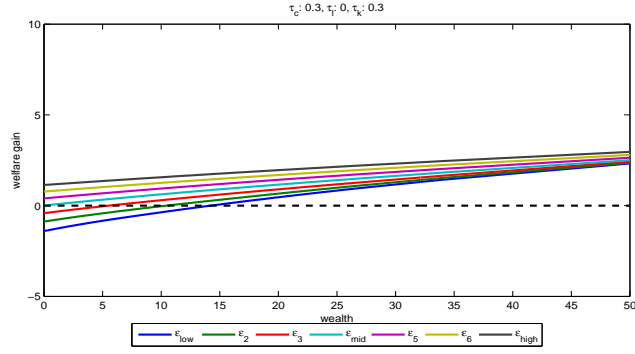
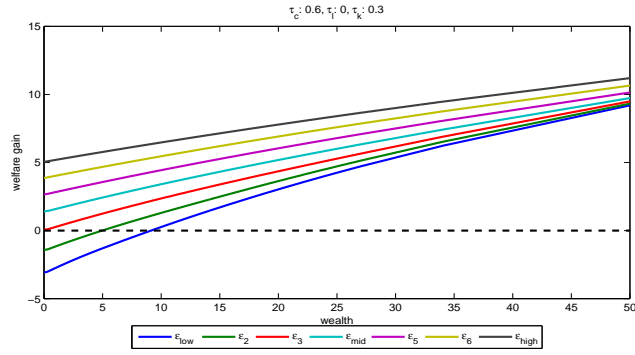


Figure 4: Preferences over consumption taxes

(a) $\tau_c = 0.30$



(b) $\tau_c = 0.60$



(c) $\tau_c = 1.20$

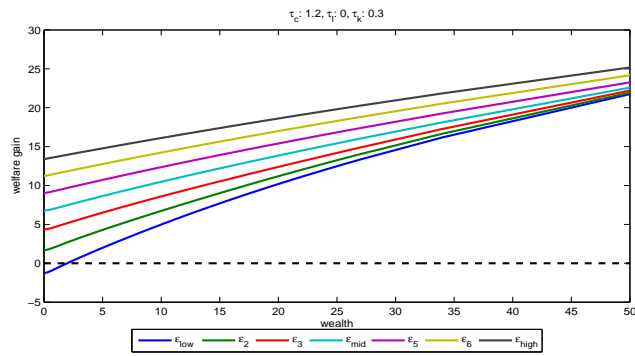
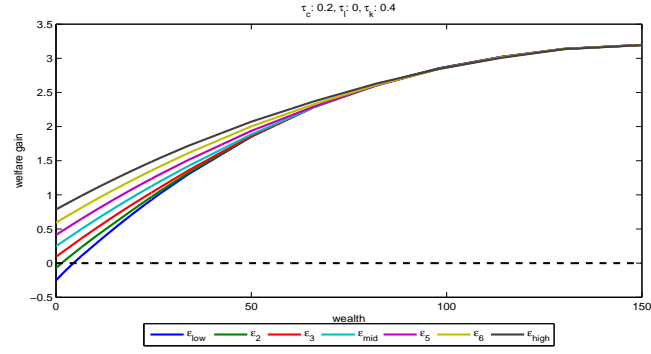
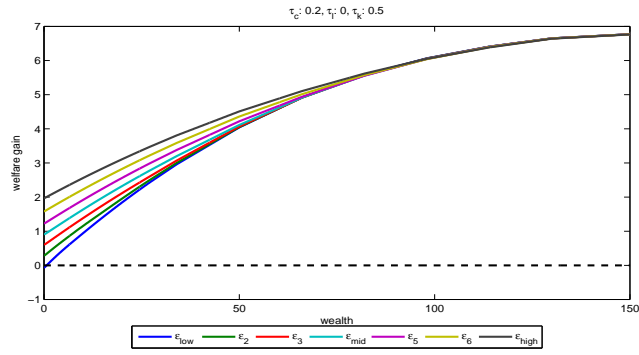


Figure 5: Preferences over capital income taxes

(a) $\tau_k = 0.40$



(b) $\tau_k = 0.50$



(c) $\tau_k = 0.70$

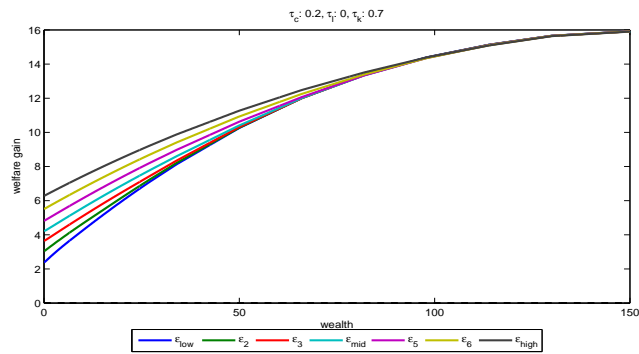
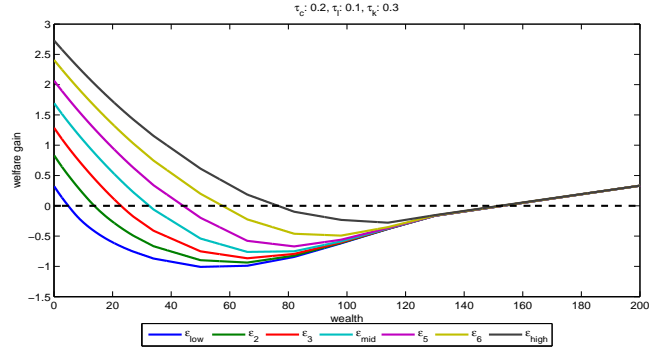
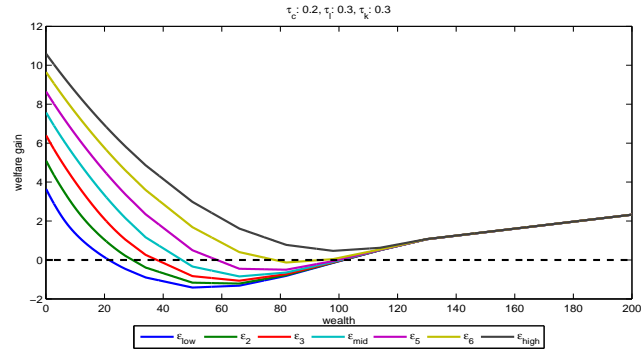


Figure 6: Preferences over labor income taxes

(a) $\tau_l = 0.10$



(b) $\tau_l = 0.30$



(c) $\tau_l = 0.60$

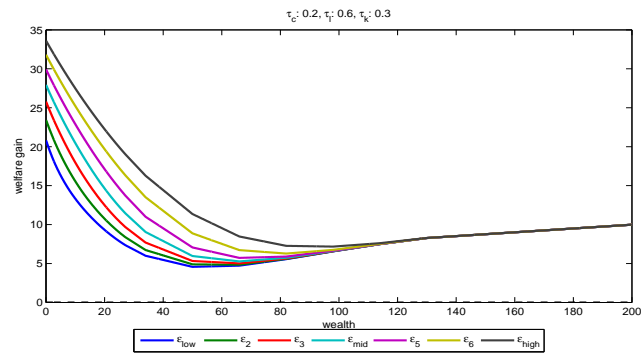
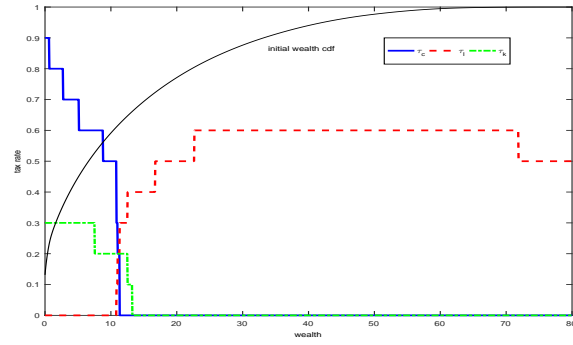
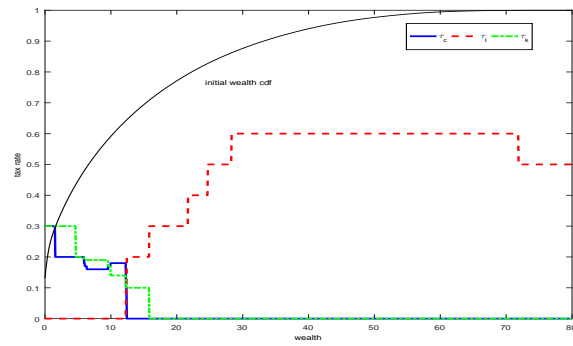


Figure 7: Most preferred tax combination by productivity and wealth

(a) Lowest ε



(b) Median ε



(c) Highest ε

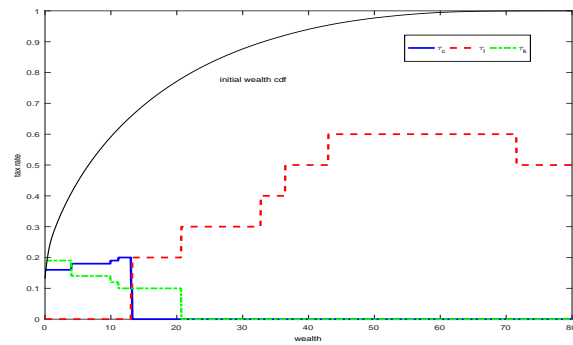
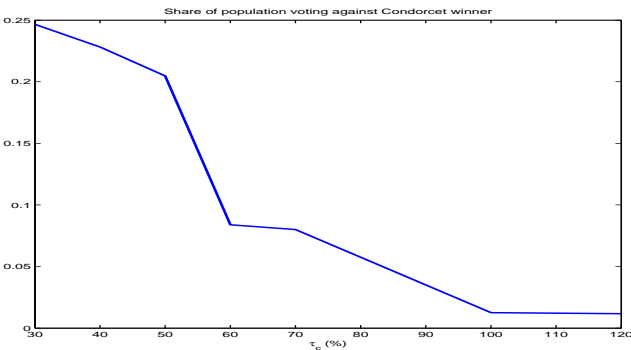
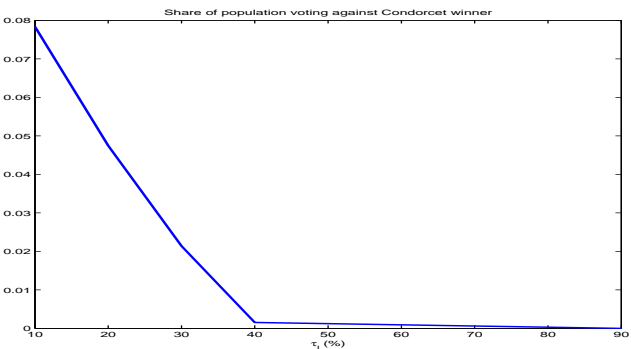


Figure 8: Share of votes for alternative policy with higher . . .

(a) consumption tax



(b) labor income tax



(c) capital income tax

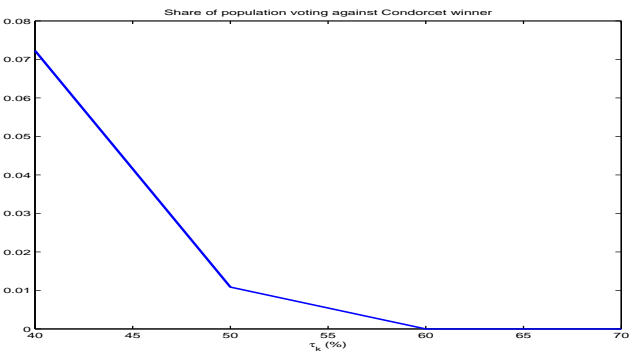


Figure 9: Transition path under Condorcet winner policy

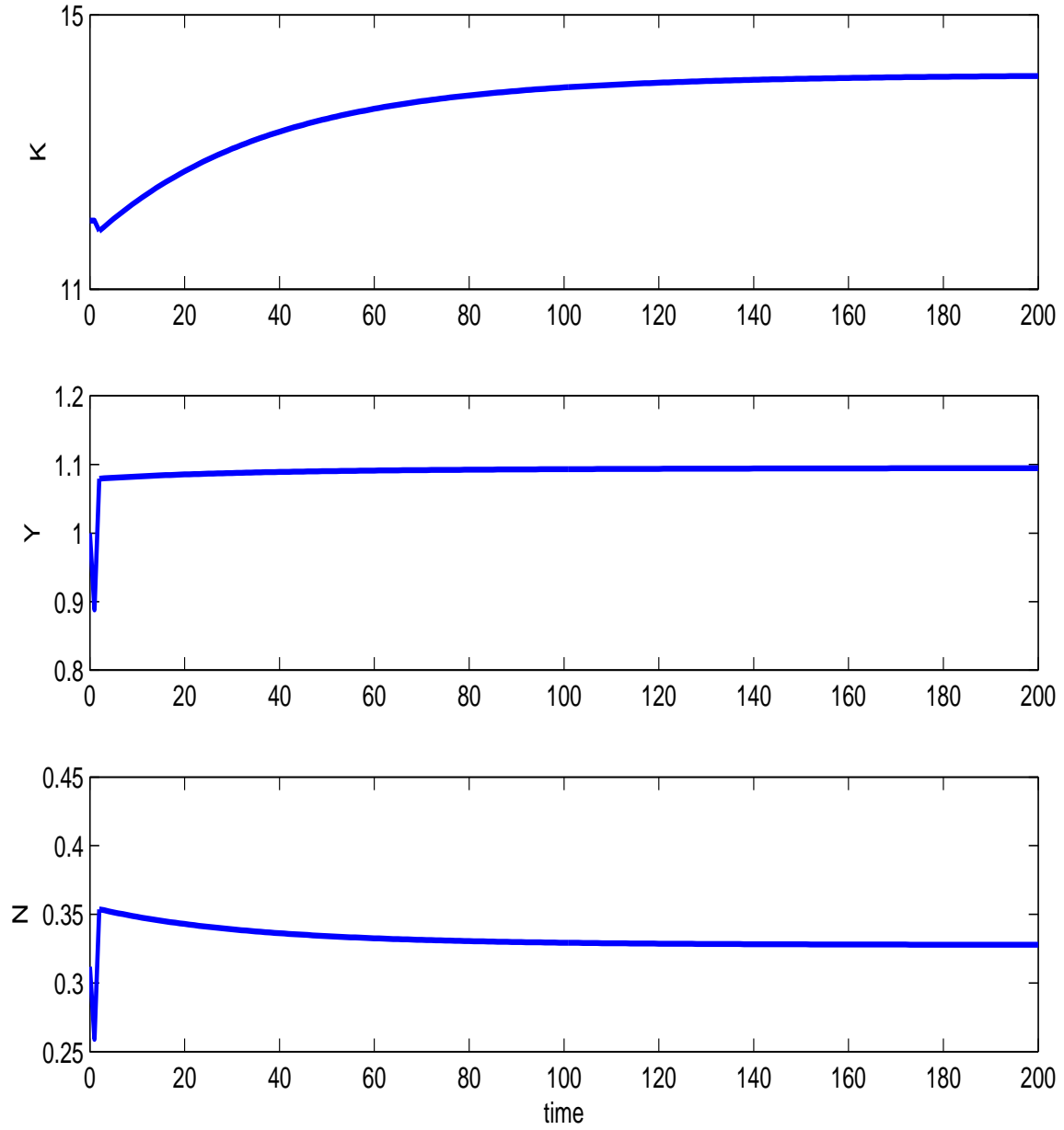


Figure 10: Transition path under Condorcet winner policy

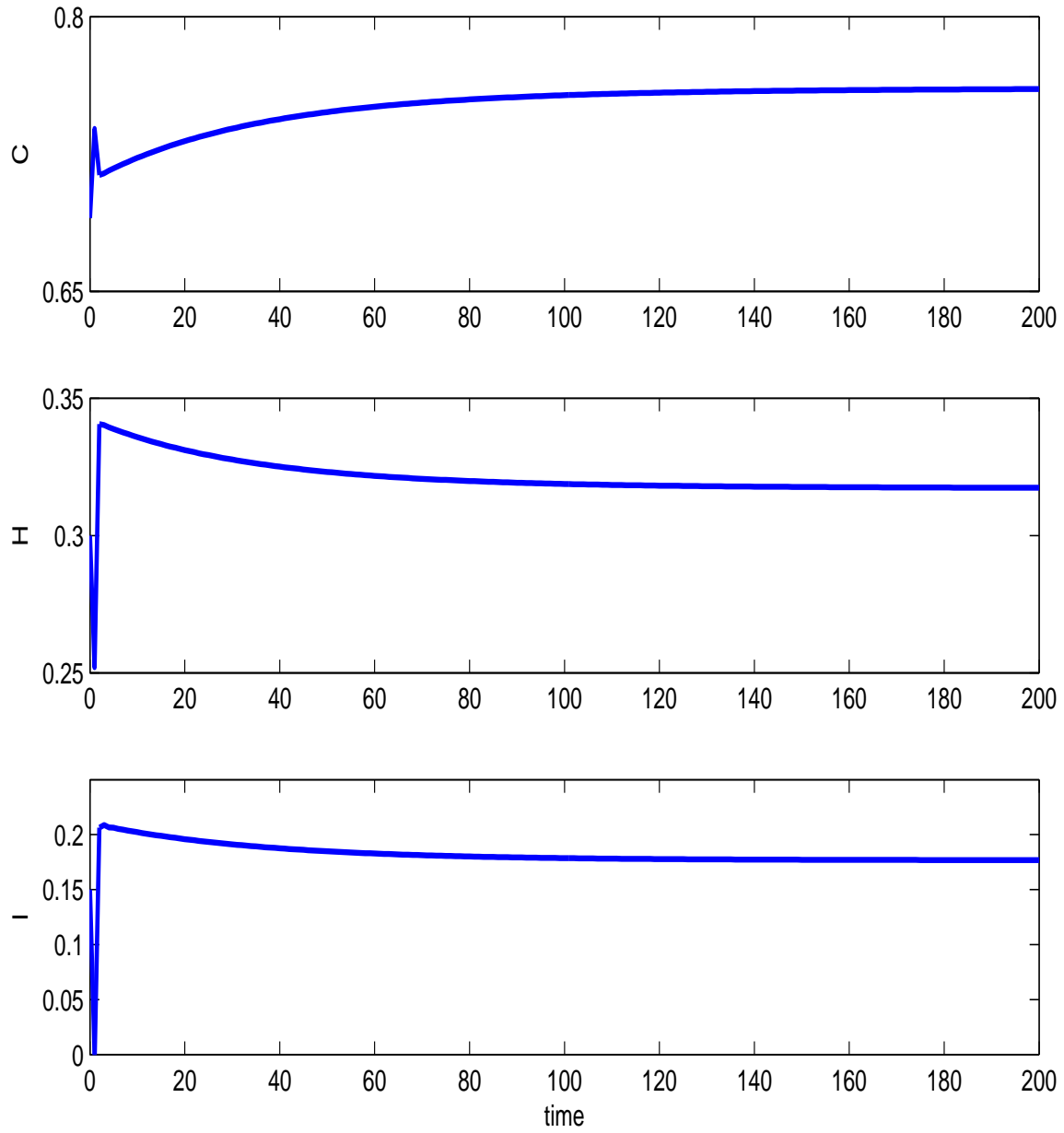


Figure 11: Transition path under Condorcet winner policy

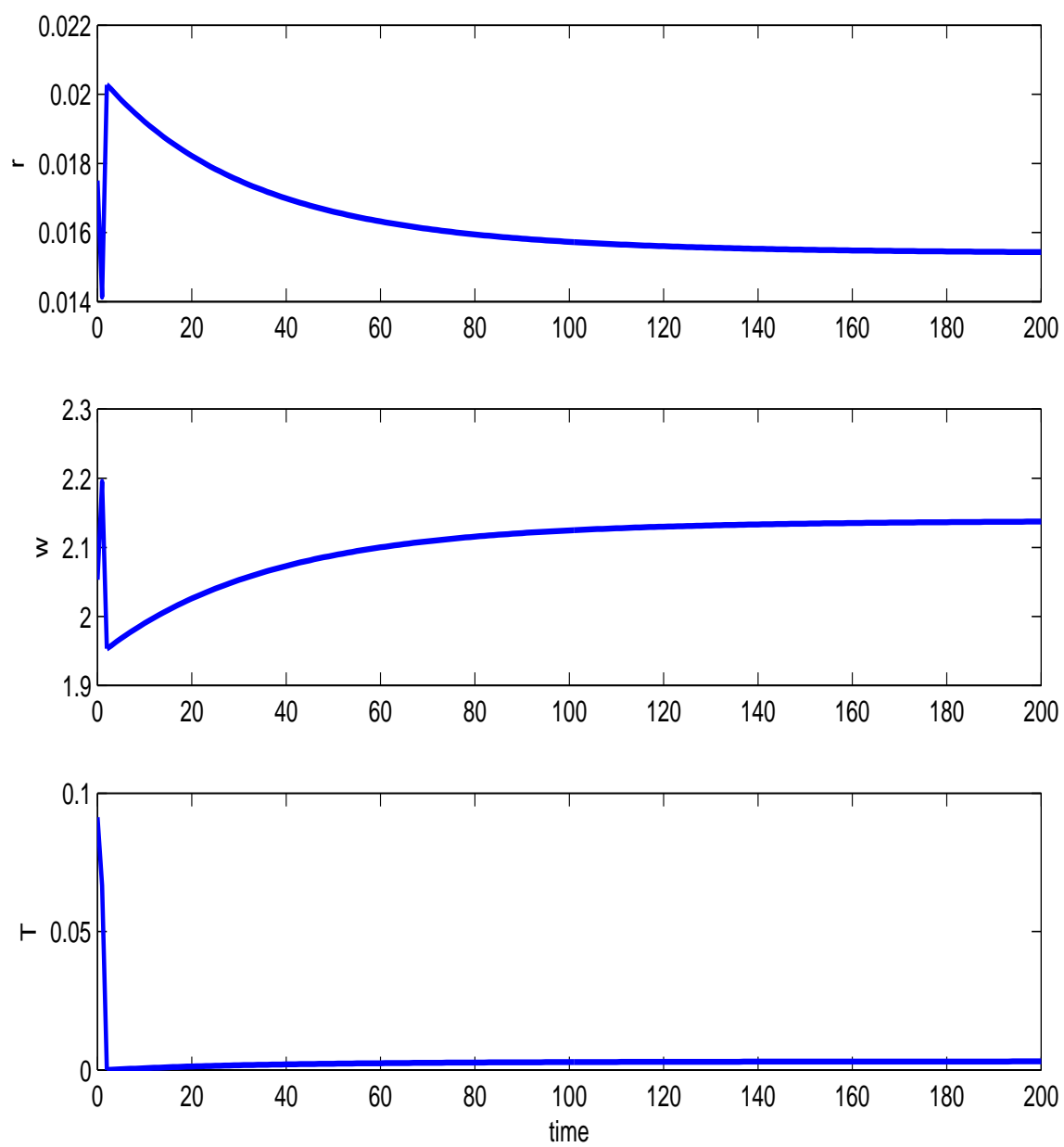


Figure 12: Wealth Lorenz Curve

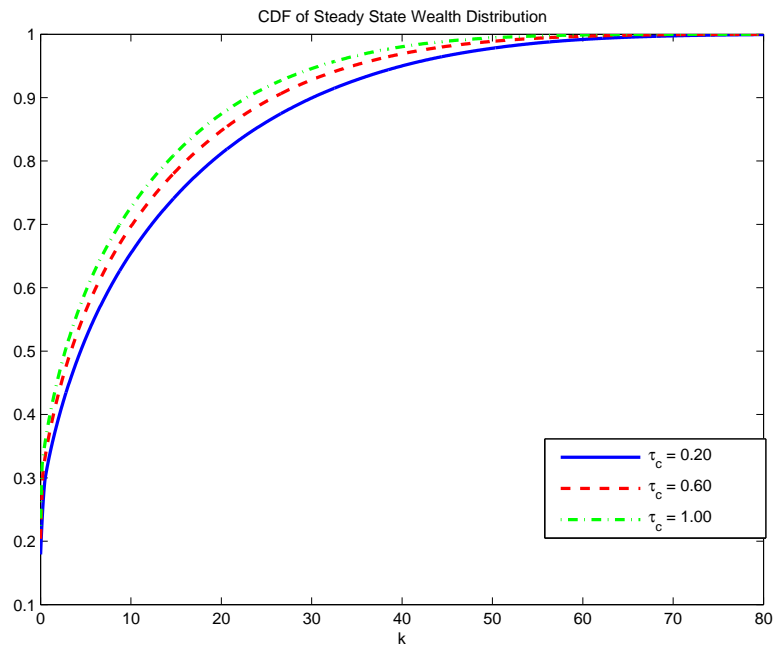
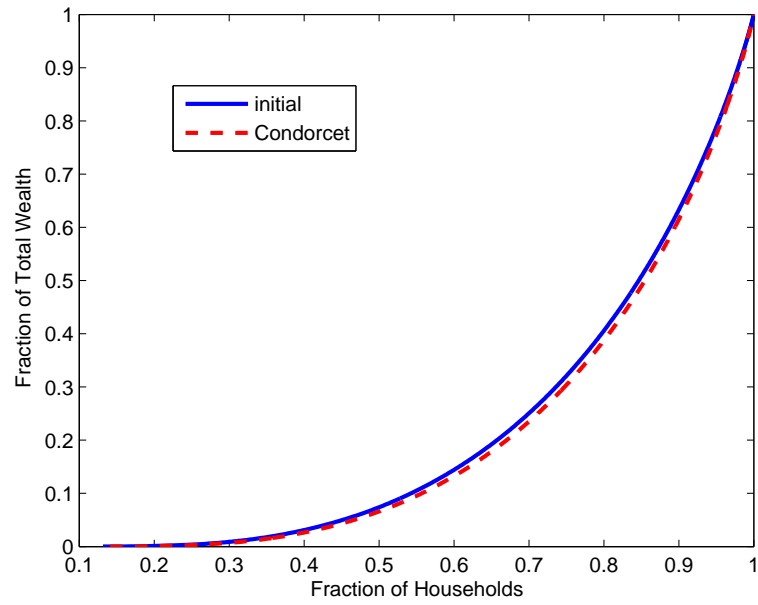


Figure 13: CDF of wealth by τ_c

Figure 14: CDF of wealth by τ_k

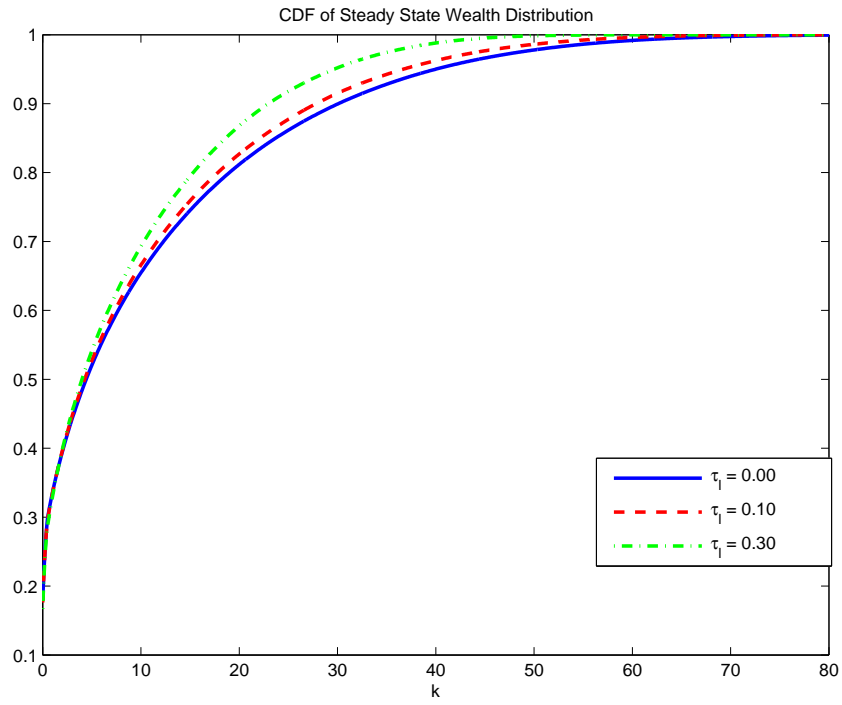
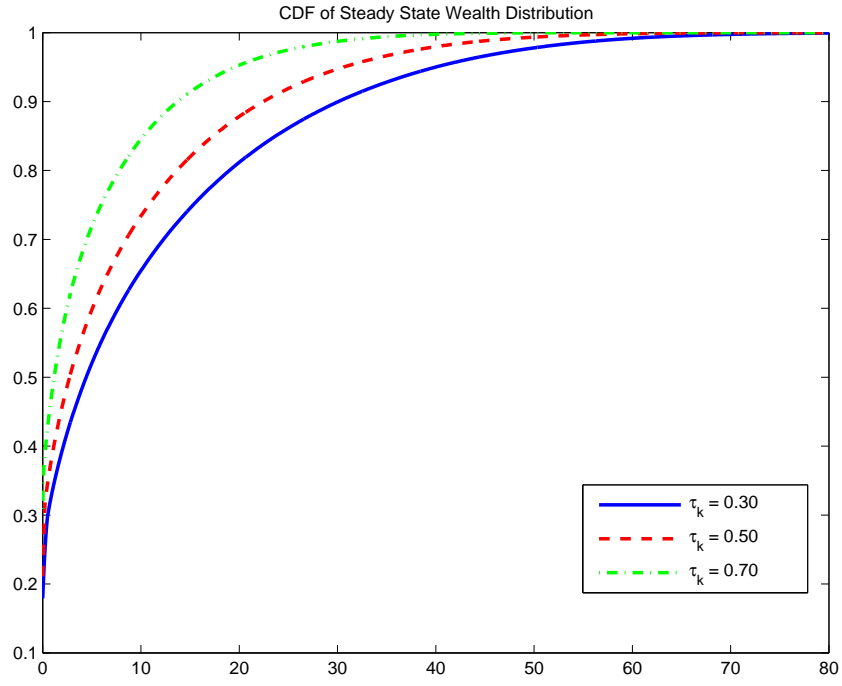


Figure 15: CDF of wealth by τ_l

Figure 16: CDF of votes under wealth-weighting

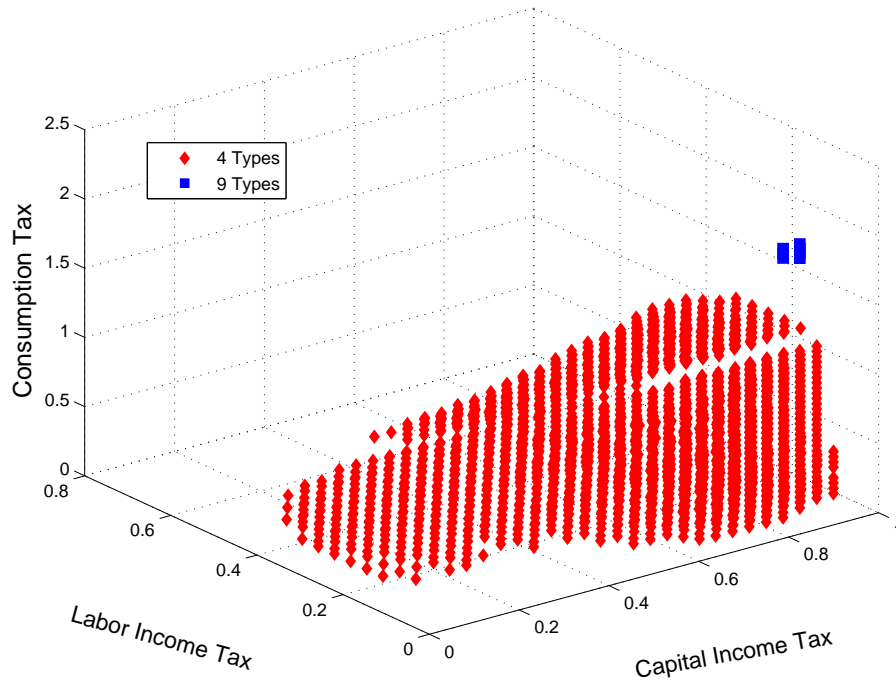
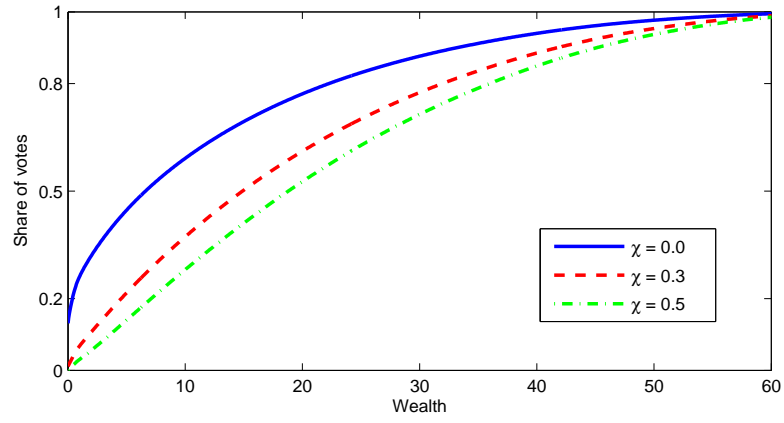


Figure 17: Uncovered set in Dolmas model

Table 1: Calibrated Parameters

Parameter	Value
β	0.986
Z	0.862
θ	2.203
T	0.091

Data	1	2	3	4	5
1	0.90	0.07	0.02	0.01	0.00
2	0.27	0.34	0.30	0.06	0.02
3	0.09	0.14	0.45	0.25	0.06
4	0.05	0.06	0.15	0.51	0.23
5	0.05	0.05	0.06	0.17	0.68
Model	1	2	3	4	5
1	0.58	0.20	0.13	0.06	0.03
2	0.16	0.35	0.21	0.20	0.07
3	0.11	0.15	0.39	0.16	0.18
4	0.08	0.19	0.09	0.41	0.22
5	0.07	0.09	0.18	0.17	0.49

Table 2: Earnings Mobility

Table 3: Wealth Mobility

Data	1	2	3	4	5
1	0.63	0.26	0.07	0.03	0.02
2	0.27	0.45	0.17	0.07	0.03
3	0.07	0.22	0.45	0.20	0.06
4	0.03	0.05	0.26	0.45	0.21
5	0.01	0.03	0.05	0.25	0.67
Model	1	2	3	4	5
1	0.76	0.23	0.01	0.00	0.00
2	0.24	0.55	0.21	0.01	0.00
3	0.00	0.22	0.61	0.17	0.00
4	0.00	0.00	0.17	0.72	0.11
5	0.00	0.00	0.00	0.11	0.89