Taxation and Public Goods under Federalism
Exit, voice, and federal redistribution

Debra Hevenstone

January 16, 2008

1 Introduction

Many countries including Australia, Austria, Canada, Germany, Italy, Spain, the United States, and Switzerland (among others) are federal governments with at least partially devolved tax systems. Inequalities in taxation and consequent public service provision can motivate citizens' residential choices and are often smoothed out across localities using national equalization schemes. The computational model simulates the interplay between households' residential decisions and local residents' influence on their governments' tax policies, finding emergent properties including economic residential segregation and unequal tax rates favoring the wealthy. Next, the model implements a federal equalization scheme showing that equalization can actually encourage more disparities by inducing wealthy areas to reduce their tax rates further. Results find that under the equalization scheme tested here, the federal redistribution scheme can only efficiently addresses about one-third of the inequalities in tax revenues across sub-governments. The model is validated using Swiss cantonal tax data although the results have implications for any federal system using an equalization scheme.

1.1 Introduction

This article unites several classic streams of social science research in an attempt to answer an important two public policy questions: First, do devolved tax systems encourage economic residential segregation, and second, if they do, do central governments' equalization schemes effectively combat that inequality? The first related area of the social science literature surrounds Tiebout's 1958 article which argued that public goods can be efficiently allocated in a devolved government in which municipalities compete, mimicking a competitive market. The second related area is the discussion about policy devolution and the "race to the bottom", and finally, the literature spawned by Hirschman's "Exit Voice and Loyalty", proposing a framework for how individuals can influence local policy, is closely related (Brandeis, 1933; Tiebout, 1956; Hirschman, 1970).

Tiebout's seminal 1956 article spawned an entire field of research, and his name has become synonymous with the idea that the combination of devolved government and residential mobility can encourage the optimal production of public goods. This idea has been influential not only in the academic field (Tiebout's 1956 article is cited by more than 2000 articles in the web of science database alone) but also in the public policy arena, where the devolution of federal policy through grants is often justified using Tiebout's argument. The idea is that the smaller governments offer different baskets of public goods and taxes and individuals purchase that basket of goods by moving there. Assuming that jurisdictions have some optimal size, that households are fully mobile and willing to move to maximize their utility from public goods, and that households have full information about all jurisdictions, an optimal provision of public goods will be reached at equilibrium. This model has been used to explain a wide variety of phenomenon including levels of general public goods provision, residential choice, residential segregation (Kessler and Lulfesmann, 2005; Epple et al., 2001), the optimal level of locally controlled redistribution (Epple and Romer, 1991), and the ideal provision of various individual public goods.

There are several major critiques of the Tiebout hypothesis. The first is the strong assumptions (Bewely, 1981), which Tiebout himself noted. The weakest assumptions are residents' perfect mobility and perfect information, both of which are far from real world conditions. Empirically households are reluctant to move, particularly in Europe compared to the United States (Bentivogli and Pagano, 1999) and households are not able to assess their utility in every possible jurisdiction. The second critique is that residential choice and public good provision is seldom focused on a broad basket of public goods as Tiebout suggests. Most people have similar tastes, and more importantly real estate prices depend on *other* people's tastes, so migration is often focused on the public services that most households want (such as school quality). People do not migrate seeking those public services targeting the poor (police, social services, public transit, etc). Consequently, those with financial resources cluster in communities able to offer schools and low tax rates, while the poor cluster in municipalities with higher tax rates providing police and social services (Orfield, 1997). Further, jurisdictions are able to distort the market and compete for more attractive high income households using tools like zoning laws to keep low-income individuals out. While the Tiebout hypothesis is a parsimonious and powerful theory, the real world the assumptions don't apply, public goods preferences correlate with income creating poor areas, and jurisdictions have preferences and tactics that promote segregation. In this model we simulate a mechanism like that proposed by Tiebout, allowing households to move based on their tax and public policy preferences.

Another social science classic, Albert O Hirschman's "Exit Voice and Loyalty" can be considered an extension to Tiebout's hypothesis. Hirschman argued that members can change the organizations and institutions (including governments) they belong to through two means: exit and voice. Exit corresponds to Tiebout's proposal in which the individual can move out of the municipality. When an individual is tied to the organization because community identity or brand name loyalty individuals use "voice" to change the organization from within. Exit and voice are replacements for one another with reactions to dissatisfaction lying somewhere on a continuum between the two, the exact balance depending on the circumstances. For example, an individual constrained from leaving a group will use voice, while in an oppressive group without free speech they will use exit. In this simulation households who do not exit are able to use voice, voting on their policy referendums.

Perhaps on of the most tested theories in political science, is whether the competition among smaller governments leads to a "race to the bottom" instead of the optimal provision of public services suggested by Tiebout. Jurisdictional competition can lead the the dismantling of state services that attract poor residents (such as social services) or that repel wealthy ones (such as high taxes). There are numerous empirical tests of this theory such as Peterson and Rom (1990) and Schram and Beer (1998) who test whether American states offer less generous welfare benefits in response to neighboring states' levels of benefits, hoping that the poor from neighboring states will not be attracted into their state. The same phenomenon is hypothesized with respect to environmental regulation or tax rates where jurisdictions lower taxes in an attempt to attract wealthy residents (Nechyba, 1997b). As more states follow this strategy, taxes and public services deteriorate. Given a devolved federal system, the first question is whether there is a race to the bottom and second how can the central government combat it while maintaining the federal structure (Breuille and Bobo, 2007)? This theory, like the Tiebout hypothesis, assumes households are very willing to move. While there has been some evidence that poor people migrated from Chicago to Milwaukee in search of lower benefits (DeParle, 2004) and that in Switzerland high income people migrate to low tax cantons (Feld, 2000), there is no case of a complete "races to the bottom" [1].

While these theories focus on the ideal case where local governments have complete control over their policies, in practice central governments mitigate the inequalities created by devolution though financial equalization and regulation. There are many different aspects of designing centralized redistribution schemes. The schemes can vary based on their caps, whether they are horizontal (money from small government to small government) or vertical (from the central government to the small governments, how it estimates and includes the small government's tax capacity, how it estimates and includes the cost of providing services in the local government, whether the money is

[1] Households fail to move despite substantial incentives. In 2004 a family of three in Alabama received $215 in public assistance compared to 709 in Vermont. In the Swiss canton Zug, a couple earning 200,000 CHF/ year only pays 8.3% in cantonal income taxes compared to 13.37, 14.4, and 15.47% in the neighboring cantons of Zürich, Aargau, and Luzern, over a 14,000 CHF saving.

transferred simply as cash or through earmarked grants, and perhaps most importantly, the marginal revenue local governments gain for tax increases under the scheme. A full description of the various types of intergovernmental transfers is beyond the scope of this paper and I refer the reader to (Boadway and Shah, 2007) or (OECD, 2007) for a thorough description of the types of programs and their incentive structures. The main goal of these equalization schemes is to balance local autonomy with the provision of equal levels of public goods for all citizens.

Unfortunately central equalization schemes can have create incentives for local governments to act strategically, increasing or decreasing their taxes (depending on the formula) or shifting public services such that the local government offers an inefficient basket of public goods (OECD, 2007). To reduce these distortions central redistribution formulae should: construct formulae based on an aggregation of all locally collected taxes; use transparent formulae; avoid 100% marginal federal taxation on locally collected revenues; measure tax capacity based on the local area's ability to collect taxes, not their actual tax revenues; avoid matching grants which might distort local preferences; use formulae not vulnerable to fluctuations that inhibit the local government's ability to project revenues; and finally attempt to construct formulae that are less susceptible to political influences. Even under the best of circumstances the equalization schemes can be create unanticipated distorting incentives. This paper tests a simple equalization scheme designed incorporating all of these criteria in an agent based model. The model is tuned using Swiss cantonal tax data although the model construction is representative of any federal devolved system.

2 Agent based models versus traditional models

The analysis public goods and taxes in a context of residential mobility is widely studied. The bulk of the literature is in the field of economics and use mathematical models to determine the equilibrium levels of segregation, whether governments set tax rates in response to their citizens (versus to attract future residents), and the natural variation in tax policy. Usually these models use linear forms of taxes (Kessler and Lulfesmann, 2005; Epple et al., 2001; Nechyba, 1997a) because the functional form is analytically tractable. However, this assumption is unrealistic particularly when we consider income tax, one of the biggest taxes in most western countries. The residential dynamics of these models are particularly sensitive to this assumption since the wealthiest are subject to higher tax rates under a progressive system and are more likely to move in response to tax rates (Feld, 2000). Translating the typical tax table into a smooth function, tax rates generally follow this form:

$$t_i = S(1 - e^{-ky_i}) \tag{1}$$

,where t_i is the tax rate for individual i, y_i is their income, and S is the maximum tax rate, and k is the phase-in rate.

The problem with this more realistic functional form is that it leads to intractable optimization problems. For example, assuming a Cobb Douglas utility function where individuals earn utility from public goods ($p = \frac{1}{n}\sum y_i * s(1-e^{-ky_i})$, from private consumption c,and from housing h, we have the already complicated utility optimization problem of

$$U_j = (\frac{1}{n_i}\sum(y_i S(1-e^{-ky_i})))^{\alpha} * (h_j)^{\sigma} * (c_j)^{\gamma} \tag{2}$$

s.t. the income constraint:

$$y_j = y_j S(1-e^{-ky_j}) + h_j + c_j$$

where,

$$\alpha + \sigma + \gamma < 1$$

This does yield 5 equations and 5 unknowns- so theoretically we should be able to find individuals' optimal tax parameters in terms of their incomes and preference rates, though this system of equations is seemingly insolvable (for me and mathematica). For your notes the Lagrangian is:

$$\mathcal{L} = (\frac{S}{n_i}\sum(y_i(1-e^{-ky_i})))^{\alpha} * (h_j)^{\sigma} * (c_j)^{\gamma} - uSy_j + uSy_j e^{-ky_j} - uh_j - uc_j + uy_j$$

yielding:

$$\begin{aligned}
\tfrac{d\mathcal{L}}{dS}:&\quad \alpha h_j^{\sigma} c_j^{\gamma}\left(\tfrac{1}{n}\sum_{i=1}^{i=n}(y_i - y_i e^{-ky_i})\right)^{\alpha-1} - uy_j + uy_j e^{-ky_j} &&= 0\\
\tfrac{d\mathcal{L}}{dk}:&\quad \alpha h_j^{\sigma} c_j^{\gamma}\left(\tfrac{S}{n}\sum_{i=1}^{i=n}(y_i^2 e^{-ky_i})\right)^{\alpha-1} - uSy_j^2 e^{-ky_j} &&= 0\\
\tfrac{d\mathcal{L}}{dh}:&\quad \sigma h_j^{\sigma-1} c_j^{\gamma}\left(\tfrac{S}{n}\sum_{i=1}^{i=n}(y_i - y_i e^{-ky_i})\right)^{\alpha} - u &&= 0\\
\tfrac{d\mathcal{L}}{dc}:&\quad \gamma h_j^{\sigma} c_j^{\gamma-1}\left(\tfrac{S}{n}\sum_{i=1}^{i=n}(y_i - y_i e^{(-ky_i)})\right)^{\alpha} - u &&= 0\\
\tfrac{d\mathcal{L}}{du}:&\quad -Sy_j + Sy_j e^{-ky_j} - h - c + y_j &&= 0
\end{aligned}$$

While some simplification is possible, the optimal tax levels S and k for individual j are not derivable. However, we can use this functional form effectively in a simulation, finding the same results that economists find using this mathematical approach, namely the distribution of optimal tax rates and the equilibrium level of residential segregation.

3 The Swiss example

The Swiss taxation system is one case of a devolved tax system. It is a particularly good case to base a model on as it has several characteristics that are easy to apply in a simple model. First, the over 44% of the income tax is collected at the cantonal level with municipal taxes (39.4%) consisting of add-ons that maintain the same functional form as the cantonal tax, leaving only 16.4% determined at the central federal level. Second, there is significant variation in these tax rates as illustrated in figure 1. Furthermore, the cantons are small, so that there is a low cost for households to move. For example, within commuting distance to the city of Zürich, are the cantons of: Zug, Zürich, Aargau, Soloturn, Luzern, Bern, Basel Stadt, Basel, Schwyz, Glarus, and St. Gallen. Given the small size of cantons, several authors have found empirical evidence of tax-based migration in Switzerland (Feld, 2000; Schmidheiny, 2006) [2] Finally, the Swiss federal government uses a direct tax redistribution scheme that can be modeled. In sum, a parsimonious and tractable simulation of tax competition and central government equalization can crafted that is representative of the Swiss case.[3] Because statistical offices publish tax tables rather than functions, we must do a little manipulation to derive the Swiss cantons' tax formulae. The Swiss tax code layers municipal, federal, and cantonal income tax rates, with cantons and federal having their own deduction schemes [4]. Rather than publishing the various deductions, rules, and rates, the Swiss statistical office publishes the income tax rates for several typical cases in each canton's capitolEIdgenossenschaft (2007). Figure 1 illustrates the tax rates by canton for a couple with no kids. Zürich, with moderate tax rates, is highlighted because it is the largest city. Zug, the lowest line with outstandingly low tax rates, is highlighted because it is also the wealthiest canton with 71,733 Swiss francs (CHF) per capita in national income. The capital with the highest rates is Delsberg, located in the canton Jura which has the lowest per capita income in Switzerland (33,616). In the first picture Sarnen is also highlighted because it has regressive tax rates and in the second picture Geneva because of its extremely progressive tax rates (up to an income of 45,000 CHF a couple pays less than one percent in income tax.)

Fitting the tax formula $t = S_i(1 - e^{ky}$ to these curves, we find approximate S and k parameters for each canton, as illustrated in figure 2. The distribution of these real-life tax rates is used to tune the simulation which is in turn used to answer three questions. First, if governments set their tax rates in response to their residents' preferences and residents move in response to tax rates, will unequal tax rates and residential segregation evolve? Second, along the continuum of exit and voice- are households more willing to move when they are dissatisfied or to vote? And finally, what is the optimal level of

[2] In contrast, in the US significant tax benefits can usually only be had by moving significant distances, or by those wealthy enough to claim residency based on a second home in a lower-tax state. In a few regions of the US there are municipal income taxes creating the same incentive (New York City (to 3.6%), Washington DC (to 7.9% & doubling as state tax), and Philadelphia (to 4%).)

[3] A more detailed description of the Swiss equalization formula is available in (Dafflon, 2004).

[4] deduction definitions, though not rates, were standardized across cantons in 2007.

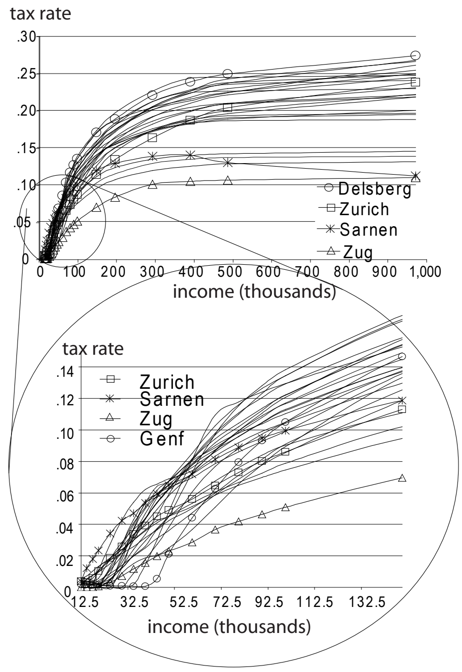


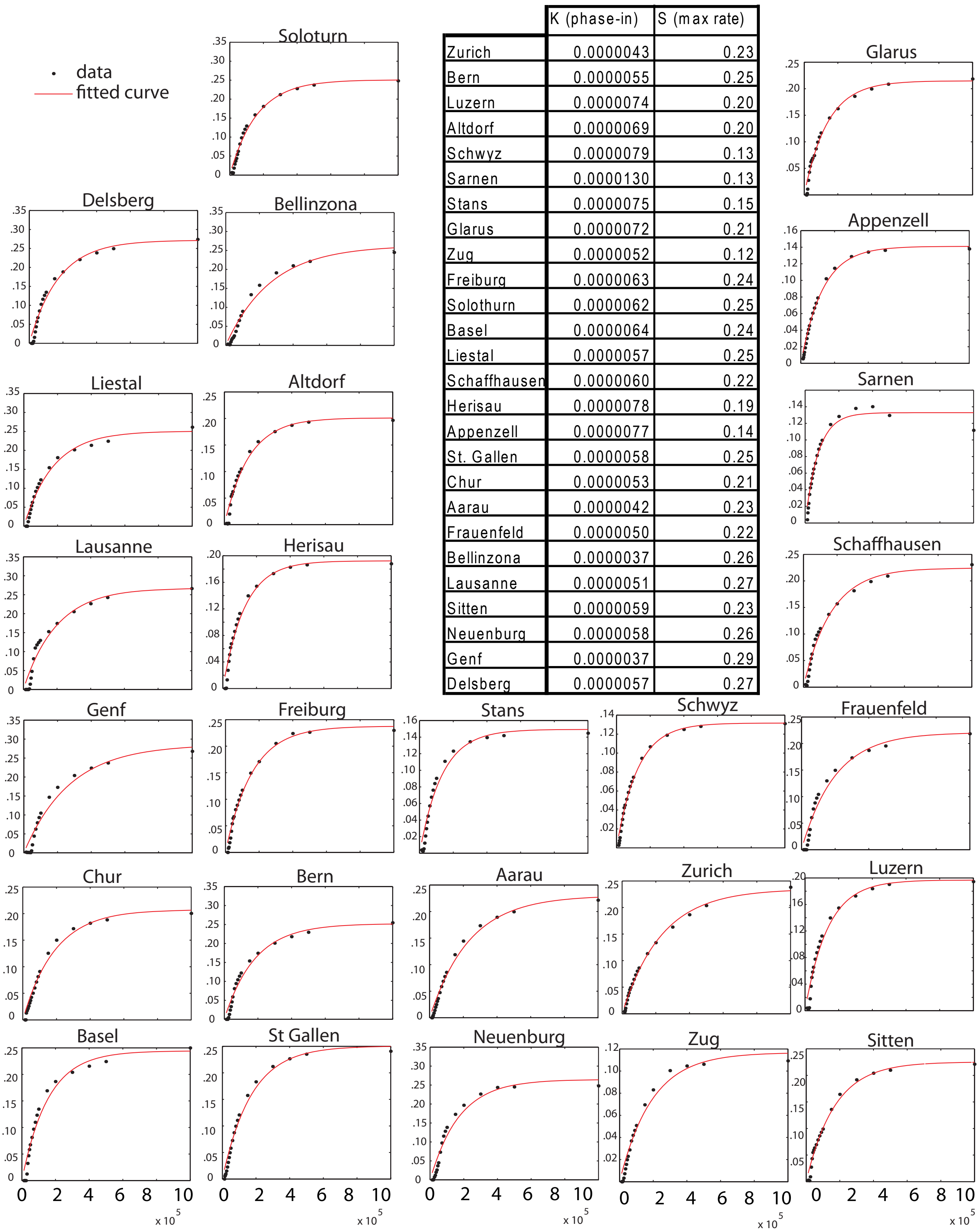
Figure 1: Cantonal Taxes for a couple with no kids

central government equalization?

4 Method

The agent based simulation models the interplay between local public goods provision and residential mobility. In this model, households move when there are a significant number of better housing options. When households do not move, they can vote on whether they want more public goods and higher taxes (or less public goods and lower taxes). Residents cannot vote for the unrealistic political rhetoric of lower taxes and more public services. Governments respond to voters by increasing, decreasing, or maintaining the same level of taxes and then the whole algorithm repeats.

The program's algorithm is illustrated in figure 3. The model takes place on a 100 by 100 torus split into 16 jurisdictions with 625 housing units per jurisdiction. Some percentage of housing units are vacant while the others are randomly assigned households. (The vacancy rate influences how quickly the model moves towards equilibrium, though not the equilibrium itself.) Occupied parcels have a starting value equal to one third of their occupant's income while empty lots have a starting value equal to the average of the neighbors' prices. The sixteen jurisdictions start with randomly assigned tax parameters: S (the maximum income tax) is assigned from a normal distribution with



	K (phase-in)	S (max rate)
Zurich	0.0000043	0.23
Bern	0.0000055	0.25
Luzern	0.0000074	0.20
Altdorf	0.0000069	0.20
Schwyz	0.0000079	0.13
Sarnen	0.0000130	0.13
Stans	0.0000075	0.15
Glarus	0.0000072	0.21
Zug	0.0000052	0.12
Freiburg	0.0000063	0.24
Solothurn	0.0000062	0.25
Basel	0.0000064	0.24
Liestal	0.0000057	0.25
Schaffhausen	0.0000060	0.22
Herisau	0.0000078	0.19
Appenzell	0.0000077	0.14
St. Gallen	0.0000058	0.25
Chur	0.0000053	0.21
Aarau	0.0000042	0.23
Frauenfeld	0.0000050	0.22
Bellinzona	0.0000037	0.26
Lausanne	0.0000051	0.27
Sitten	0.0000059	0.23
Neuenburg	0.0000058	0.26
Genf	0.0000037	0.29
Delsberg	0.0000057	0.27

Figure 2: Backing Out Cantonal Tax Formulae

$\bar{X}_S = .35$ and $\sigma_S = .076$ and k (the phase-in rate) is also assigned from a normal distribution where ,$\bar{X}_k = -.0000225$ and $\sigma_k = .0000000001$. Households have Cobb Douglas utility functions, $U = p^\alpha * h^\sigma * c^\gamma$ with randomly assigned preferences for public goods (α), housing consumption(σ), and private consumption (γ), which sum to one. Households are randomly assigned incomes pulling from a normal distribution with mean 11 and standard deviation of 1 and setting income equal to e to the power of that draw. This creates a skewed income distribution with the mean household income at 60,000, the 95th percentile at 425,000 and the 5th percentile at 22,026.

Once households, parcels, and jurisdictions are generated, jurisdictions calculate their public service provision. They do not have economies of scale or different efficiencies in producing public goods [5]. Households produce public goods as the sum of tax revenue divided over the number of households in their jurisdiction, $p = \frac{1}{n} * \sum_{i=0}^{i=n_t} S_t * (1 - e^{-k_t y_i}) * y_i$.

After public services are calculated, households have the opportunity to move. The results presented here give households the chance to move in a fixed order although random with and without replacement was also tested. Households make their decision by calculating their private goods consumption (subtracting taxes and housing from income) and then calculate their current utility function. Once households know their current utilities, they randomly select ten vacant lots and calculate their potential utilities if they moved. The potential utility takes the current level of public services as given, assuming the single household's tax contribution will not change the average level of public service provision. This utility is set to zero if the cost of housing and taxes exceed the household's income. Finally, if roughly proportion (**E**) of the sampled properties offer higher utilities than the household currently has, the household moves to the best of those options. Just after moving, households are ineligible to vote for one round, until they establish residency.

Once all households have had the chance to move, it's election season. Every jurisdiction proposes a higher and a lower tax scenario to its voters, adjusting the S and k parameters with a shift from a uniform distribution from -.02 to .01 for S and 0 to .000001 for -k. Households calculate their prospective utilities under the proposals and vote -1 for a tax reduction or 1 for a tax hike. The jurisdiction then averages the votes and if it is above (below) the positive (negative) threshold, they increase (decrease) taxes. Otherwise, taxes stay constant. Once voting is over, the whole cycle starts again following the sequence: move, vote, set taxes.

[5] Other papers use a link function where jurisdictions have diminishing returns, or where jurisdictions can become unproductive at some capacity (Kessler and Lulfesmann, 2005)

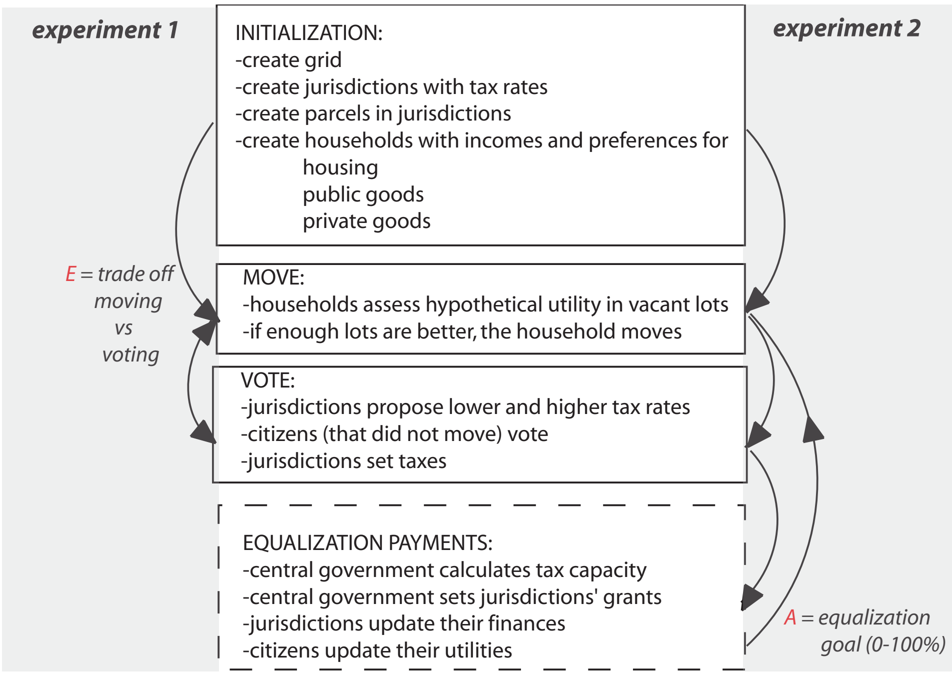


Figure 3: Program algorithm

4.1 Experiments

In the first experiment, we vary the parameter **E**, controlling how willing individuals are to move and give up their vote. This variable is theoretically the balance between exit and voice for each individual. In the second experiment, an algorithm that redistributes revenues across jurisdictions according to their tax capacity is introduced. In this experiment the central government seeks to redistribute between 0 and 100% of the inequality in tax capacity. All combinations of parameters run in each experiment was run in ten separate simulations for a length of 1,000 steps for each simulation. The presented results come from the last time step and represent the averages across the 10 simulations. While data are presented by jurisdiction, jurisdictions with the same rank by income were averaged across the ten simulations with the same settings.

The first experiment has two important components. First, it shows that even when governments simply respond to their constituents (not competing for wealthier households) tax rates in wealthier areas decline and households segregate by income[6]. Second, we find a value for the E parameter (the trade-off between exit and voice) that generates realistic tax rates. The resulting setting for E suggests that individuals are not entirely mobile, which most economic studies assume. The second experiment tests how much the central government can redistribute revenues without generating perverse incentives

[6] (Epple et al., 2001) shows that the dynamics change as we introduce strategic governments.

for jurisdictions.

4.1.1 Experiment One: Segregation, Exit, and Voice

The first experiment varied the parameter, E, controlling the tradeoff between exit and voice. Specifically, E measures the proportion of viewed housing options that must be better than a household's current residence, to motivate a household to move. When the household moves, it loses its right to vote for one round, until it has established residency. E was varied from .05 to .5 meaning that they move if between 5 or 50% of the sampled alternative homes had to offer a better utility to instigate a move.

In figure 4, each line on the graph indicates a jurisdiction with three jurisdictions highlighted- the poorest, the wealthiest, and the average of the middle 2 jurisdictions. Each panel shows the results for the average across ten simulations holding the parameter, E (the exit-voice parameter) constant. The x axis shows income and the y axis illustrates the tax rate. As such, the graphic shows taxes as a function of income, as the prior graphs of the Swiss cantons did. The shape of the curve is governed by the two tax variables (S and k) which are the maximum tax bracket rate and the phase-in rate. The dots along the tax curve indicate the mean income in the jurisdiction with that tax schedule.

In the first panel we see the tax rates picked by the richest, poorest, and middle income jurisdictions holding the parameter E at a mean of .05 where households move even when only one hosing alternative is better. Under this setting, the poorest jurisdiction has the highest overall tax rate, and the mean income is on the portion of the curve that is steeply sloped, meaning that within the jurisdiction individuals pay varying tax rates, or along a progressive schedule. In contrast, the richest jurisdiction has a lower rate for the highest tax bracket, but a quicker phase-in rate, meaning that their tax is more similar to a flat tax, with most residents paying the same rate. Across jurisdictions the tax rate is quite regressive with the average earner in the rich jurisdiction (255,000 CHF) paying about the same rate as the average household in the poor jurisdiction (50,000 CHF). While the first panel depicts the expected result (the wealthiest jurisdiction has slightly lower taxes) this pattern is only stable through the first four panels where households are relatively willing to move. At middle levels of mobility, from $E = .3$, the model goes through a transition with unstable tax rates (with wide variation from run to run) where the richest jurisdictions maintain the lowest overall tax rates but with the transition rates varying such that the average income household in the middle jurisdictions pay higher taxes than those in the rich. By the time households are heavily substituting voting for moving (E = .35), the richest jurisdictions have much lower and flatter taxes while the poor and middle income jurisdictions bring their maximum tax brackets up to 50%, although the tax rate for the average resident stays at approximately the same tax rate of 20%. In other words, with less of a threat that the wealthiest will leave, the poorer jurisdictions settles on high and steep tax rates that extract revenue from the wealthy.

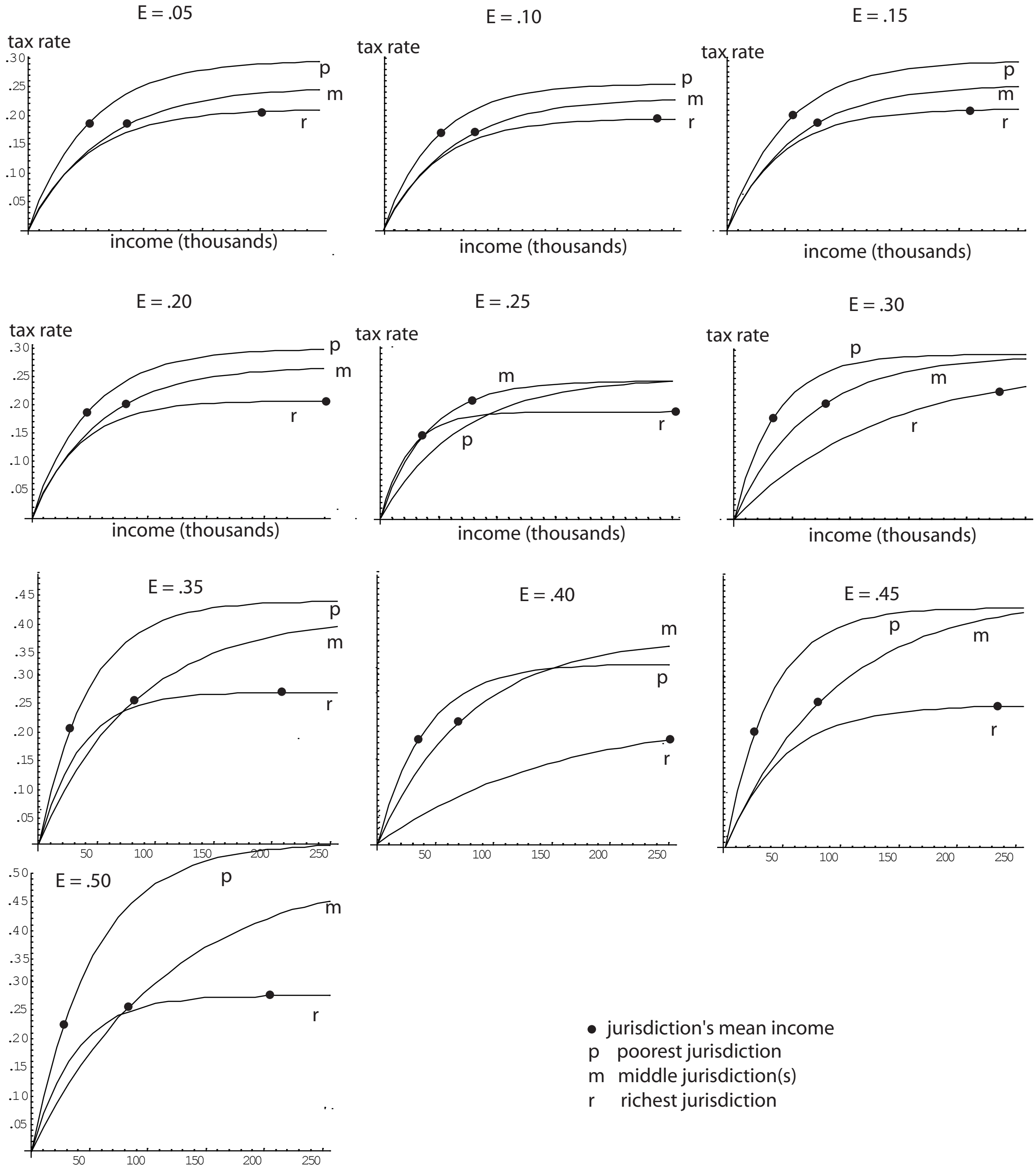


Figure 4: Tax schedules by jurisdiction, varying the exit parameter (E)

Figure 5, shows a broader overview of the simulations, also illustrating a period of unstable dynamics around E = .3. In the first panel we see the gini coefficient across the jurisdictions' mean incomes, showing economic segregation in all the simulations which is possibly *higher* when residents are less willing to move. The second two panels show the standard deviations for the tax parameters across jurisdictions and suggest that when households are more likely to move than vote, tax rates vary less across jurisdictions. As with the gini coefficient, there is a transition at E = .3 after which tax rates become very unequal. While the overall trend in these three figures suggests that less mobility and more voting lead to both more segregation and unequal tax rates, around the transition point of E =.3, there is a small return towards less segregation and equal tax rates.

Why does a shift towards voting and away from moving increase segregation and unequal tax rates? Perhaps with high mobility, the composition of communities changes to quickly for residents to vote consistently. However, when the communities remain the same, the government has more time to adapt to residents' preferences, increasing inequality. This is particularly likely since the model limits tax proposals to small deviances from current policy. While low exit and high voice might engender inequality and unfair taxes, the last plate on figure 5 shows that it also increases utility when citizens can change their government's policies to suit their preferences. While the increase in voice is perhaps more efficient, it is also more unequal. Depending on one's belief that social welfare is dependent on the average person's utility or the poorest person's utility, increasing mobility could be considered either positive or negative.

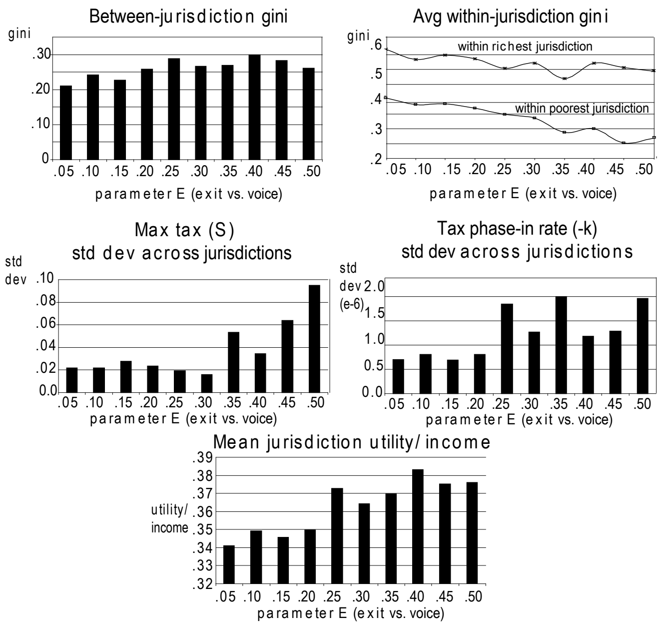


Figure 5: Aggregate statistics, varying the exit parameter (E)

Figure **??** shows that the gini coefficient within jurisdictions decreases as people are less likely to move, again suggesting increasing segregation. Because incomes are skewed, there is always more intra-jurisdictional inequality for the richest jurisdictions. Again, around E = .3 trends become more unstable.

Returning to the Swiss tax rates illustrated in figure 2, the real-world mean value of -k is -.062$*e^{-4}$ while in the simulations the overall average was -.039, a somewhat slower phase in rate than in real life. The only place the simulation is very close to this phase in rate is when E = .2 , a balance between exit and voice, just before the critical point of $E = .3$. The standard deviation of K for the Swiss cantons is 1.84 $*e^{-6}$, which around the same as it is for K under the higher E simulations as illustrated in figure 5. The variation in phase in rates across cantons is similar to in the simulation. The absolute level of phase in rates in real life is similar to the simulations with a moderate trade off between exit and voice. The maximum tax rate (S)in real life is lower than in the simulations. However, adding in the federal rates, they are about equivalent. The average tax rate across cantonal capitols was 22% of income with a standard deviation of .0479 while the model produces similar maximum tax brackets until E reaches the critical value of .3. While the simulation tax rates match the cantonal rates at low E's (or the cantonal plus federal at higher E's) the simulation's inequality across jurisdictions only is as high as in real life at higher E's. In sum, the actual tax levels we see in real life match those simulations where individuals have a moderate trade off between exit and voice, though the level of inequality in real life is closer to the model's inequality just at the threshold, around E = .25 or .3.

The debate about federalism and the race to the bottom focuses on a scenario where the local governments compete, lowering their taxes to attract richer residents. However in the model we show that this same pattern emerges naturally when governments are simply responsive to their residents and when residents are mobile. Assuming, that the central government values a progressive tax regime with the richer generally paying higher taxes, there are several policy approaches a federal government can take to alleviate this situation. This brings us to experiment two, where we test one possible equalization strategy.

5 Experiment Two: Tax Equalization

Many countries (Canada, Switzerland, Australia, Belgium, and Germany) with devolved tax systems have complex formulae for redistributing revenue from those jurisdictions with high tax capacities and/or low expenditure needs to those with low tax capacities and/or high expenditure needs. While the American system primarily relies on matching grants for individual programs with matching weights based on states' capacities, other countries, like Switzerland directly redistribute revenue either vertically from the federal government to the jurisdictions for horizontally between jurisdictions. Theoretically this redistribution calculates the tax capacity of the jurisdiction and compares it with the average tax capacity across jurisdictions, charging those with more capacity and reimbursing those with less. In this simulation we use this simple formula:

$$G_i = (R_i^* - R_i) * A * P \tag{3}$$

,where G is the federal grant, R_i^* is the amount of revenue collected per capita for the aggregate of jurisdictions using the average of all jurisdictional tax rates, R_i is the per capita revenue collected using jurisdiction i's population with the average tax rates, P is the population of the jurisdiction,and A is some value between 0 and 1, indicating the amount of federal redistribution. When $A = 1$, the federal government insures perfect equality and when A=0 the central government does not redistribute any revenue. This is normally adjusted so that $\sum G_i$ is equal to the federal budget for redistribution. In reality, the formulae is much more complicated as revenue comes from various taxes, and formulae often weight the different tax revenues, relying more heavily on those taxes with more consistent revenues. Formulae can also incorporate macroeconomic measures such as the regional gross product. Finally, many formulae include adjustments for the cost of providing public services or the level of need in the region.

In Switzerland there are no less than 10 formulae for redistribution (Dafflon, 2004) with all relying theoretically on the following formula (retaining Dafflon's notation):

$$E_i = \frac{1}{5}[(1.5 * NIC_i) + (\frac{100}{B_i} * 100) + 1.5(\frac{T_i + \sum T_{im}}{H_i} * \frac{100}{B_i}) + .5 * (\frac{U_i^p}{U_i} + \frac{H_i}{km^2})] \tag{4}$$

This equation includes tax-based measures of capacity, macroeconomic measures of capacity, and two geographic measures designed to measure the difficulty of providing services in the individual jurisdictions. The 1/5 in the front of the equation standardizes the weights 1.5, 1, 1.5, .5, and .5 to one and the indices i and m indicate cantons and municipalities, respectively. The first component, weighted by the 1.5 (a 30% weight) is the national income per canton (NIC_i) per capita (H_i). The second measure, weighted by 1 (a 20% weight), is the inverse of the global index of tax burden B_i. B_i is calculated by considering the tax rate for 5 cantonal taxes: individual income, individual wealth, institution profit, institution capital, and motor vehicle taxes. The categories of taxes are then weighted by their relative importance for revenue in the given canton, and are then standardized such that the average tax burden is 100. The third element, weighted by 1.5 (30%) is the sum of tax revenue in the canton and municipalities per capita, adjusted by the tax burden. Another way to think about this is that it is the per capita revenue times a maximum of 2 for the lowest tax burden district or times .75 for the highest tax burden district. The final two elements weighted .3 each (or 10%) are $\frac{U_i^p}{U_i}$ the ratio of agricultural land in the plain region to total agricultural land and $\frac{H_i}{km^2}$ the population density. The last two elements benefit low population and mountainous regions, assuming that it is more difficult to deliver services in these areas. The final index is then adjusted such that the lowest jurisdiction has a score of 30 (Valais), meaning the highest jurisdiction, Zug, has a score of 218. Individual program formulae expand on this equation, adding a benefit for post-industrial areas as well so that both rural mountainous areas and cities benefit. This formula has been the basis for tax redistribution since 1959 and was not changed until 2007. For more details on the old system and the new reforms the reader can directly consult (Dafflon, 2004) for a thorough description of the Swiss tax equalization scheme or (Boadway and Shah, 2007) for a discussion of the theoretical issues behind these formulae.

In the model we use the simple grant calculation illustrated in equation 3, testing how jurisdictions set their tax rates in response to the central government's decision to pursue no equalization (A= 0) to full redistribution (A= 1). Using this formula, when high-capacity jurisdictions tax their residents at very low rates, the central government demands more grant money than the jurisdiction has collected. When this happens, the jurisdiction gives the central government 100% of its revenues and provides no public services, but does not collect any additional revenue from its citizens. The equalization algorithm is appended to the earlier model directly after jurisdictions select their tax rates. The equalization algorithm first calculates the negative grants, charging high-capacity jurisdictions and updating their residents' utilities. Next, it redistributes that money among those jurisdictions with positive grants. However, since positive and negative grants do not balance when the central government bankrupts a rich jurisdiction, the jurisdictions with positive grants receive only $\frac{G_i^+}{\sum(G_i^+)} * G_i^-$, weighting their grant by their need as a proportion of all need across jurisdictions. Finally, the residents in grant-receiving jurisdictions update their utilities.

5.1 Findings

As in the prior experiment, rich jurisdictions charge lower tax rates and phase in taxes quicker, as shown in figure 6, which shows the average tax rates for the poorest, middle, and richest jurisdictions under various settings of the redistribution parameter, A. The diagram clearly shows that as soon as the central government implements any equalization scheme, rich jurisdictions' maximum tax rates decline to about 10% and when redistribution reaches 20 percent, the phase in rate for rich jurisdictions accelerates, with the phase in to the top bracket complete for those with incomes over 100,000 per year. The poorest jurisdictions have the highest tax rates and the most progressive tax schedules and the middle-income jurisdictions have just slightly less progressive schedules such that the average resident in a middle-income and a poor jurisdiction pay the same tax rate.

In terms of the inequality across tax rates, as illustrated in the fourth and fifth panels of figure 7, the inequality in maximum tax rates and phase-in rates grows rapidly as the central government redistributes revenue. Perhaps one of the most interesting outcomes of the model is illustrated in the top panel of figure 7. Here we see that in the scenario with no redistribution, as before, wealthy jurisdictions provided much more public services. However, as the rich jurisdictions' public funds are redistributed to the other jurisdictions, they reduce their tax collection and their public service provision. When central government redistributes one-third of the inequality in revenue, the richest jurisdictions provide the same public service as the others. Any attempted increase in equalization beyond this point undermines redistribution as the rich jurisdictions simply drop out and stop taxing their citizens. The model currently runs 1,000 steps, to the point that migration is relatively stable and that public service provision in rich areas reaches zero. At step 1,000 these rich jurisdictions are still collecting taxes from their population on behalf of the central government. Given that these payments are a pure loss for all individuals in the community, and projecting forward, we can assume that eventually these jurisdictions will pay no taxes at all.

If we consider the jurisdiction as well off as its poorest individual (often referred to as a Rawlsian definition of poverty), without revenue equalization it is better to be poor in a rich area. This situation rapidly reverses as rich jurisdictions stop collecting taxes and the poor people in these areas stop receiving public services. (Note that the income of the poorest household in a jurisdiction are relatively constant around 3,000/ year.) Again, the optimal level of equalization is around A = .3.

Oddly, federal equalization schemes do not reduce residential segregation in the model, and might actually increase it, as shown by the third panel in figure 7. This happens for two reasons: first, the housing pricing mechanism assures that areas that start a bit richer by pure chance, will continue to have more expensive housing, excluding the poor. Second, federal equalization creates incentives for rich jurisdictions to reduce taxes. This lower tax rate compensates families for their loss in public services by allowing them more consumption. However for poor households with lower incomes, it is better to stay in the less wealthy districts where they still have access to public services.

Finally, the results suggest that on average the equalization scheme reduces average utility. The last panel of figure 7, shows that equalization beyond A = .3 yields decreasing average returns on utility per unit of income. This is clearly because in a free market the wealthy would prefer to consume more public and less private goods. However, given the threat of their public revenue being taken by the central government, they substitute for less satisfying private consumption. In addition, the poor in wealthy jurisdictions are hurt by the lack of public goods.

6 Conclusion

There were three main findings from this experiment. First, the interplay between residential mobility and simple government responsiveness was found sufficient to create unequal tax systems where rich jurisdictions have low taxes and poor jurisdictions have high taxes. Normally this dynamic is thought to result from strategic jurisdictions

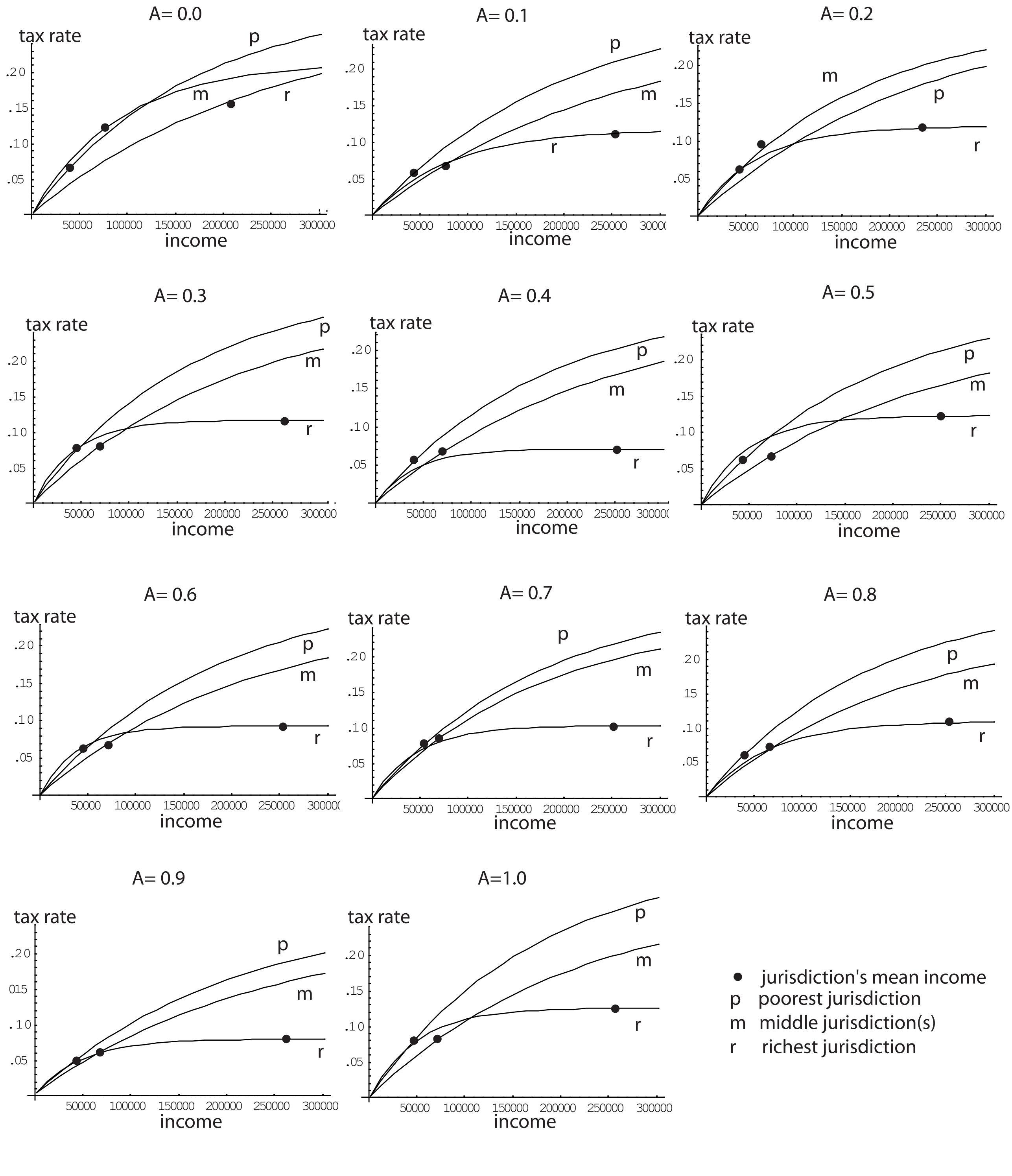


Figure 6: Tax schedules by jurisdiction, varying the federal redistribution parameter (A)

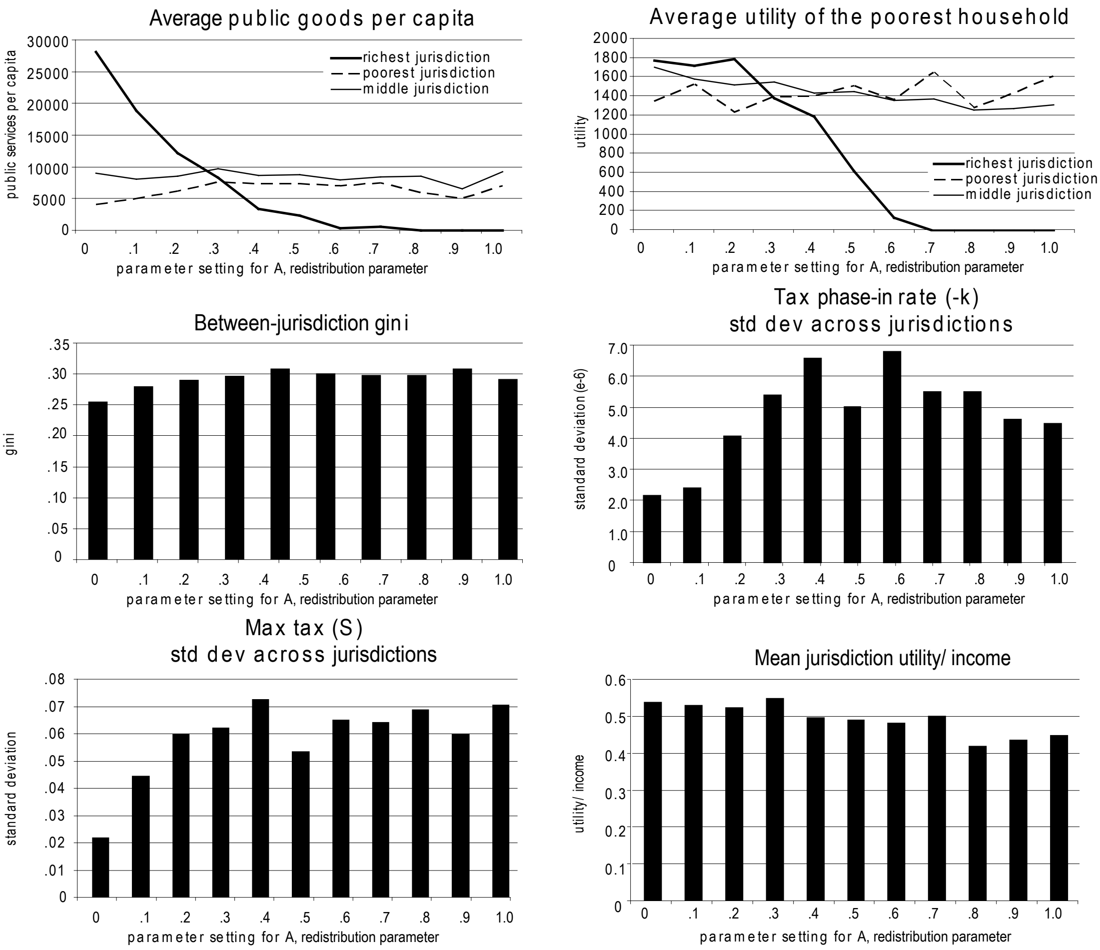


Figure 7: Aggregate statistics, varying the redistribution parameter (A)

competing for residents. Second, using Swiss data on cantonal tax rates, it seems that the model generates more realistic tax rates when individuals are moderately reluctant to move. Third, a centralized system of revenue equalization can create perverse incentives for the local jurisdictions, encouraging rich areas to reduce their tax rates even further than they would under a devolved tax system without equalization. Results suggest that redistributing more than a third of the inequality in tax revenues is infeasible. Of course, these results do not signify that equalizing public services for residents is inviable, only that if a central government wants a truly equal system of taxes and public service provision, perhaps they need a centralized one.

7 Bibliography

References

Bentivogli, C and P Pagano. 1999. "Regional disparities and labour mobility: The Euro-11 versus the USA." *Labor* 13:737–760.

Bewely, Truman F. 1981. "A Critique of Tiebout's Theory of Public Expenditures." *Econometrica* 49:713–740.

Boadway, Robin and Anwar Shah. 2007. "Intergovernmental Fiscal Transfers: Princeiples and Practice." Technical report, World Bank.

Brandeis, Justice Louis. 1933. "Ligget Co. v. Lee."

Breuille, Mari-Laure and Robert J Gary Bobo. 2007. "Sharing budgetary austerity under free mobility and asymmetric information: An optimal regulation approach to fiscal federalism." *Journal of Public Economics* 91:1177–1196.

Dafflon, Bernard. 2004. "Federal-Cantonal Equalisation in Switzerland: An overview of the present system and reform in progress." Technical Report 356, Centre d'Etudes en Economie du Secteur Public.

DeParle, Jason. 2004. *American Dream.* Penguin.

EIdgenossenschaft, Schweizerische. 2007. "Steuerbelastung in der Schweiz:Kantonshauptorte- Kantonsziffern 2006." Technical report.

Epple, Dennis and Thomas Romer. 1991. "Mobility and Redistribution." *The Journal of Political Economy* 99:828–858.

Epple, Dennis, Thomas Romer, and Holger Sieg. 2001. "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis." *Econometrica* 69:1437–1465.

Feld, Lars P. 2000. "Tax competition and income redistribution: An empirical analysis for Switzerland." *Public Choice* 105:125.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty*. Harvard University Press.

Kessler, Anke S. and Christoph Lulfesmann. 2005. "Tiebout and redistribution in a model of residential and political choice." *Journal of Public Economics* 89:501–528.

Nechyba, Thomas J. 1997a. "Existence of equillibrium and stratification in local and hierarchical Tiebout economies with property taxes and voting." *Economic Theory* 10:277–304.

Nechyba, Thomas J. 1997b. "Local Property and State Income Taxes: The Role of Interjurisdictional Competition and Collusion." *The Journal of Political Economy* 105:351–384.

OECD. 2007. "Fiscal Equalisation in OECD Countries." Technical report.

Orfield, Myron. 1997. *Metropolitics*. Brookings Institution Press.

Peterson, Paul E and Mark C Rom. 1990. *Welfare Magnets: A new case for a national standard*. Brookings Institution Press.

Schmidheiny, Kurt. 2006. "Income segregation and local progressive taxation: Empirical evidence from Switzerland." *Journal of Public Economics* 90:429–458.

Schram, Sanford S and Samuel H. Beer. 1998. *Welfare Reform: A race to the bottom?* Johns Hopkins University Press.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *The Journal of Political Economy* 64:416–424.

8 Appendix- Program Details

Global variables:

- Activation order: list, random w/o replacement, random w/ replacement
- Alternative Homes The number of vacant homes agents compare.
- Housing Start The startup value of housing (percent of households' income).
- Jurisdiction Can be set to any even square (in this experiment set to 16). May have effects on outcomes and should be tested before we hold it constant to test the parameters of interest.
- Mean and Variance for alpha, gamma, and sigma Sets the parameters controlling households' preferences for housing, public goods, and consumption.

- mean and variance income Set households' income.
- mean and var K and S Set the starting tax parameters for the jurisdictions.
- occupancy Sets percentage of the land parcels that are occupied.
- sizeX and sizeY The size of the grid.
- Tax Scenario Takes values 1,2, and 3 which specify whether the referendum proposes changes to S (the maximum tax bracket), k (the phase-in rate), or both.
- **mean and var E** Sets the parameter determining what percentage of alternative homes need to be better for a household to move.
- Change Threshold How distant the vote needs to be from 0 (no change) towards -1 or 1, for the jurisdiction to change taxes.
- Redistribute Sets whether the experiment includes central government intervention algorithm or not. This is equivalent to setting the next parameter to 0 or nonzero.
- **A** Sets the level (from 0 to 1) that the central government will attempt to redistribute jurisdictional tax revenue.

Instance Variables by Class

- Jurisdictions
  - id
  - S & k Tax rate parameters.
  - highS & highK Proposed tax increases.
  - lowS & lowK Proposed tax decreases.
  - pubGoods Sum of tax revenue divided across households.
  - lowPGoods Proposed public goods under the low tax alternative.
  - highPGoods Proposed public goods under the high tax alternative.
  - jurisdX and jurisdY Used to match contiguous parcels to jurisdictions.
  - parcelList The list of parcels that are in the town's jurisdiction.
  - residentList The list of households residing in the town.
- Parcels
  - id
  - price
  - x & y parcel location.

- myHousehold Household on the parcel.
- myJurisdiction Jurisdiction parcel belongs to.

- Households

  Households are the primary actors, making residential and voting decisions, and are consequently the most complicated class of objects.

  - id
  - x & y
  - income
  - E Threshold for moving.
  - color Shade of blue based on income (for GUI).
  - my public Public goods from the jurisdiction
  - myHousing Spending on housing
  - myConsumption Spending on private consumption.
  - myTax Spending on taxes.
  - α, σ, γ coefficients in Cobb Douglas utility formula for public goods, housing consumption, and private consumption respectively.
  - myUtility $U = p^{\alpha} * h^{\sigma} * c^{\gamma}$
  - vote less taxes (-1), same (0), or more taxes(1).
  - home Their parcel.
  - jurisd Their jurisdiction.
  - moved variable indicating if they just moved (0/1)

Startup

- Jurisdictions

  Jurisdictions are randomly assigned locations (x,y). They start with tax parameter S, assigned from a normal distribution with $\bar{X}_S = .35$ and $\sigma_S = .076$ yielding a maximum tax rate ranging from about .2 to .5. If a negative value is drawn, the number is resampled. The tax phase-in rate, k, is also set from a normal distribution with $\bar{X}_k = -.0000225$ and $\sigma_k = .0000000001$.

- Parcels

  Parcels assigned to all locations on the grid and is assigned a jurisdiction such that all parcels in a jurisdiction are contiguous and such that the all jurisdictions have the same number of parcels. (If an experimenter accidentally sets an odd ratio of jurisdictions to grid size, the jurisdictions around the margins of the grid are expanded to include the extra parcels.) Occupied parcels start with prices equal to to .33 of the current occupant's income and vacant parcels are set as a mean of adjacent lots.

- Households

  Households are then randomly given a housing parcel, the accompanying location, and then assigned public goods based on their jurisdiction, and figure out their consumption as the remainder of their income. Income is set by pulling x from normal distribution with default values of a mean of 11 and a standard deviation of 1, which are then used to generate $y = e^x$. This creates a skewed income distribution with the mean household income at 60,000, the 95th percentile at 425,000 and the 5th percentile at 22,026. Households start with their consumption preferences set from distributions defined by the aforementioned global parameters were $\bar{X}_\alpha$ (preference for public goods) = .15, and $\sigma_\alpha = .075$, housing preference σ has $\bar{X}_\sigma = .3$ and $\sigma_\alpha = .1$, and private consumption preference γ has $\bar{X}_\gamma = .65$ and $\sigma_\gamma = .1$. These draw are then normalized so that they sum to one. No preferences are allowed to be 0 and values are resampled if that occurs.

Dynamics

- jurisdictions calculate public services: $p = \frac{1}{n} * \sum_{i=0}^{i=n_t} S_t * (1 - e^{-k_t y_i}) * y_i$
- households calculate their tax, consumption, and utility
- households examine 10 alternative homes' utilities.
- with more than E% better alternatives, the household moves to the best one.
- households that just moved cannot vote
- jurisdictions propose alternative taxes. The model can be set to propose changes to the maximum tax bracket S, the phase-in rate k, or both. The proposed shifts range from -.2 to .2 for S and -.000001 to .000001.
- jurisdictions calculate alternative public services
- households compare current, high, and low scenarios and vote.
- jurisdictions averages the votes across households and over a threshold above/below 0 (.1 in this experiment) they adjust taxes.
- central government calculates tax capacity of jurisdictions and assigns redistributive grant
- jurisdictions with negative grants pay the central government the full grant amount or their entire budget, whichever is smaller.
- central government distributes proceeds to jurisdictions with positive grants.