© 2009

Soumyanetra Munshi

ALL RIGHTS RESERVED

ESSAYS ON DEMOCRATIZATION AND TAXATION

by

SOUMYANETRA MUNSHI

A Dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Economics

written under the direction of
Professor Tomas Sjöström
and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2009

ABSTRACT OF THE DISSERTATION

Essays on Democratization and Taxation

by SOUMYANETRA MUNSHI

Dissertation Director:

Professor Tomas Sjöström

This dissertation contains three chapters on political economy. Chapter 2 develops a game-theoretic model of democratization while chapter 3 studies women's suffrage in the United States empirically. Chapter 4 is a theoretical chapter exploring taxation with endogenous income.

In chapter 2, I propose a political-economic model of democratization. In this model, part of the electorate wants extension because this makes their favorite outcome in future, more achievable. And to increase their current probabilities of winning, parties propose extension, even if that means moving away from their favorite platforms in future. Here extension of franchise occurs under the following circumstances: an almost even distribution of partisans in the population, large rents from office and a particular party enjoying some partisan advantage among the voters. The mechanism does not explicitly incorporate redistributive aspects of franchise extension and hence is a more plausible model for instances of enfranchisement where redistributive repercussions may not have been a potent consideration, like women's suffrage.

In chapter 3, I study women's suffrage in the United States empirically. Though women's suffrage was federally mandated in the United States by the nineteenth amendment in 1920, many states had granted suffrage to women prior to that and most of them were clustered in the west. I revisit some of the popular conjectures that have been put forward to explain why these states moved first to give women the vote and offer a hypothesis of partisan competition leading to suffrage extension. I find evidence that early enfranchisement of women in the western states was driven by the intensity of competition between Republicans and Democrats, as well as by adverse female-male ratios, greater concentration of the population in urban areas and by a neighboring states adoption of women's suffrage. Also, the 'risk' of suffrage enactments was increasing over time.

In chapter 4, I study income taxation with endogenous income. In a voting-over-income-taxation game, there exists no pure strategy equilibrium when voters' incomes are exogenous. This is true even if the space of tax-schedules is restricted to be marginally progressive only. However in such a setting, with endogenous income, pure strategy equilibrium exists.

Acknowledgements

It would be an injustice to and a gross understatement of my profound respect for my advisor if I just state that I wish to express my sincerest thanks to my advisor Prof. Tomas Sjöström. Of course, I do want to thank him for his numerous invaluable suggestions helping me get around impediments and his supportive encouragement at all times. The strategies he suggested at every step during my graduate study, regarding research, teaching and other academic endeavors, were infallible, to say the least. But in every encounter that I have had with him, I came away with much more than just good modeling ideas - they were there, of course, and were very ingenious at that- but it was really his astute way of dealing with problems, in both Economics and real life that I really valued. His advice, I admit, often was not the easiest to follow because understandably, he wasn't trying to make life appear a cakewalk, by solving my problems himself. So in case of a hurdle he wasn't trying to lend a hand to help me go over it, rather he would try to make me strong and capable enough to face and overcome it - and this was both in Economics and outside of it. I am sure I have many weaknesses, personal and academic, but somehow he never made me either aware or ashamed of them. Rather he often tried, I thought, to inculcate more confidence in myself and help me look for the better things in myself. His advicing provided just the right space that I thought I needed for research, while standing beside me at all times I needed him. Prof. Sjöström once said that an advisor for Ph.D is advisor for life. He certainly will be so for me.

All my thanks also to Prof. Carolyn Moehling without whose constant support

and help the emprical chapter would never even have been started, far completed. She was extraordinarily patient with me inspite of my utter inexperience with empirical work which I'm sure must have vexed her often. I am really grateful of her careful reading of my drafts so many times. I will owe any future empirical work that I pursue to her enthusiasm and support.

My thanks also to Prof. Colin Campbell for giving me valuable suggestions regarding improving my papers. I would also like to thank Prof. Hugh Rockoff for kindling my interest in American economic history in his lively history class. My thanks also to Prof. Oriol Carbonell-Nicolau for leading me into the beautiful world of political economy and his patient guidance in all my academic pursuits. My thanks also to Prof. Richard McLean for taking so much interest in my work and pointing to possible pitfalls. My thanks to Prof. Jeffrey Ritter for finding time to read my voluminous dissertation and giving me a political scientist's perspectives on it.

Last, but not the least, I would like to express my deep love and gratitude to all the staff of the department, especially to Dorothy. Just like there could not be Hogwarts without Hagrid, there could not have been the department without Dorothy. Her constant affection and care for me made Rutgers truly a home away from home.

Dedication

To

baba, mamma

and RM

Table of Contents

Abstract ii

Acknowledgements iv

Dedication vi

1. Introduction 1

2. Enfranchisement from a Political Perspective 10

2.1. Introduction 10

2.2. The General Model 19

2.3. Subgame Equilibrium 22

2.4. Equilibrium in Period 1 Election 25

2.5. Intuitive Interpretation 27

2.6. An Example 31

2.7. Historical Evidence 40

2.8. Conclusion 49

3. Partisan Competition and Women's Suffrage in the United States 51

3.1. Introduction 51

3.2. Partisan Competition and Suffrage Extension 58

3.3. Other Hypotheses 63

3.4. Empirical Methods 68

3.5. Data 71

3.6. Analysis and Results 78

3.7. Checking for Robustness 89
3.8. Conclusion 99

4. On Existence of Equilibrium with Endogenous Income 100
4.1. Introduction 100
4.2. Notation and Definitions 104
4.3. A Benchmark: Exogenous Income 107
4.4. The Model with Endogenous Income 109
4.5. The Main Findings : Summary 136

Appendices 140
.1. Democratization-A 140
.2. Democratization-B 140
.3. Democratization-C 141
.4. Democratization-D 142
.5. Women's suffrage-A 144
.6. Taxation-A 148
.7. Taxation-B 152
.8. Taxation-C 153
.9. Taxation-D 154
References 157
Curriculum Vita 164

Chapter 1
Introduction

vidyayā amṛtam aśnute

(Knowledge leads one to immortality)

This dissertation theoretically and empirically looks at the phenomena of democratization and taxation. It contains three chapters pertaining to these issues. In two of these chapters I have used game-theoretic techniques in formulating models and solving them in trying to answer certain interesting questions in the above-mentioned areas. I also have an empirical chapter closely related to my theoretical chapter on democratization. Below I provide summaries of the three chapters:

Chapter 2: Enfranchisement from a Political Perspective

This chapter deals with extension of franchise or democratization. Various models of democratization have been proposed in the literature, like enfranchisement in response to threats of revolution (Acemoglu and Robinson 2000, 2001, 2005), threats of war (Ticchi and Vindigni 2005), split of interests among elites (Llavador and Oxoby 2005), and others. There are also different historical instances of democratization presented in support of these theories. However there seems to be instances of democratization that do not follow any of these patterns.

For one, these theories rely heavily on redistributive repercussions of extension of franchise and hence consider cases of extension from 'elites' to 'working classes'. First of all, there is very little empirical evidence that redistribution played any

major role in democratization experiences of countries. Even if it did, there are cases of democratization like women's suffrage, where redistribution would not be a major concern. My game-theoretic model is an attempt to fill in this gap.

The model captures the idea of democratization arising out of forces entirely within the economy, namely evenly-balanced partisan competition between political parties, large rents from office and a part of the enfranchised benefiting from extension of voting rights (where benefits are possibly unrelated to suffrage directly). My model fits the democratization experiences of several countries like Sweden, Chile, Italy. Also, since the mechanism of my model does not depend on redistribution, I can explain instances of suffrage extension with little or no redistributive repercussions like women's suffrage. Examples include women's suffrage in the states of the United States and countries like New Zealand and Australia.

My model is a two-period model of election where in the first period the issue is whether or not to extend franchise. And in the second period some other policy is under consideration (like prohibition, anti-child labor legislation etc). Both voters and parties have preferences over the second-period policy. Everybody is rational and forward-looking and though they do not have any direct preference over suffrage, they realize that who gets to vote potentially affects future policies for which they care.

I model two effects working in two (possibly opposing) directions for a party when it is considering extension - a 'direct effect', which means that extending franchise affects the probability of winning in this period; and an 'indirect effect', which captures the effect that franchise extension today will potentially affect future policies for which the parties care.

The 'indirect effect' is likely to be positive for the party whose next-period favorite

becomes more achievable due to extension in period 1. It is likely to be negative if extension drives outcomes away from one's favorite. The direct effect, again can go either way - some voters today would like extension because it drives tomorrow's outcome closer to their favorite while the opposite happens for others. The party proposes extension if the positive effect outweighs the negative so that the sum of the effects is positive. Here, under large rents from office and one of the parties enjoying an advantage with respect to distribution of partisanships, I get large and positive direct effects for both the parties, so that both the parties propose extension.

And part of the electorate benefiting from extension, makes the 'direct effect' positive. This is how. Under certain conditions (which include almost equally distributed partisan preferences, my notion of political competition here, and large rents from office), there is convergence of the second-period policies further from the median voter's favorite. This happens because the stronger party (the one in whose favor partisanships are distributed) is likely to propose a policy too close to its own favorite. Since rents from office are high, the other party also announces a policy nearby. This leaves a majority of the voters (like the progressive-minded people in US states) today, whose favorites are further from the policies likely to emerge tomorrow, supporting extension, if that means coming closer to their own favorites. In other words, the voters today will try and shift tomorrows electorate today because that brings the possible policies tomorrow closer to their own favorites. The other voters who were happy before, obviously do not support extension, but on the whole the parties find that extension increases their probability of winning this period (positive direct effect), when the total marginal gain in utility outweighs the total marginal loss in the current electorate.

Future Plans:

There are several ways in which the model can be made richer for a fuller representation of reality. The following are some of them: moving beyond the two-period set-up towards an infinite time horizon; letting parties have concerns for the potential rise of a third party (with the support of the newly enfranchised possibly); and considering different electoral rules and their repercussions on extension decisions.

It would also be interesting to see what kind of extension model, would fit democratization along ethnic/racial lines. For example, are there models that explain extension of voting rights to aboriginals in Australia, the black in some countries? It will also be enlightening to incorporate disenfranchisement in this framework (to explain for example, people with criminal records getting disenfranchised).

Chapter 3: Partisan Competition and Women's Suffrage in the United States

This is an empirical chapter. The starting point of this analysis (which has also been a motivation for the first theoretical work on suffrage) is that some of the states in the United States had granted suffrage to women over about thirty years prior to 1920, the year in which womens suffrage became federally mandated. Now there seems to be a discernible geographic pattern to these suffrage states these were mostly the western states (Wyoming, Colorado etc).

However the results of systematic analyses so far to explain such a geographic clustering have been most inconclusive. This chapter is an attempt to fill this gap. Here I employ event history analysis (EHA) methods to determine what factors might have contributed to granting voting rights to women in these states.

Suffrage states had certain peculiarities as far as political, social, economic and other characteristics at the state-level were concerned. For example, politically, these states seem to be most closely contested between the Republicans and the Democrats; the progressives also made headway in those states first; demographically, the population in these states was largely young-male dominated. From these and many other observations in the literature on womens suffrage, I come up with several testable hypotheses. Those pertaining to the political scenario in these states are especially interesting for me, given my theoretical democratization model (in the first chapter) where enfranchisement occurs under close balance between partisans and a part of the electorate wanting extension (like progressives and young men in the west).

In this chapter, I revisit the popular conjectures within an event history analysis framework, which is an appropriate tool for analyzing policy adoption decisions in different places over time. I also propose the hypothesis that political competition, qualified in a suitable way, can explain early suffrage in the western states significantly. With clustering of states in the west, I also look into the issue of suffrage diffusion. The idea of diffusion is that states adopt a particular policy not only as a response to its internal characteristics (like demographics and politics) but also to adoption decisions of neighboring states (where neighbors could be states sharing geographical borders or more generally, those having similar incomes or similar proportion of blacks). I complicate the model by incorporating several time-varying covariates (like sex-ratio, percentage of urban population and political competition, appropriately defined, among others) as well as those fixed over time (like geographical location of a state) and ask how each of these factors might have influenced the timing of suffrage.

I find strong evidence that early enfranchisement of women in the western states was driven by the intensity of competition between Republicans and Democrats,

as well as by adverse female-male ratios and greater concentration of the population in urban areas. Moreover, as might be expected from the geographic concentration of the suffrage states, I find evidence that suffrage adoption was strongly and positively related to whether a neighboring state had women's suffrage. Also, the risk of suffrage enactments was increasing over time foreshadowing the success of the nineteenth amendment.

Future Plans:

Now 'suffrage' or voting rights itself are multi-tiered, especially in the context of women's suffrage - it could be right of voting in school elections, municipal elections or on tax and bond issues, presidential elections, or full suffrage which means voting in all of them. ('Suffrage' in the chapter mostly includes full suffrage.) Hence I would like to estimate a competing risks model where, I analyze the question, what is the probability of a state granting a particular type of suffrage, given it had no suffrage of any kind or suffrage of some other kind? Incorporating various time-varying and time-fixed variables in this framework will also be interesting.

Alternatively, we could think of different categories of suffrage really having an underlying hierarchy (so that voting at the school level is the lowest level of suffrage while voting in presidential elections, is the highest). Hence we could also employ some kind of ordered risks model and again incorporate various variates to make the analysis richer.

In fact, an even richer environment would be to allow for disenfranchisement possibilities (suffrage rescinded in the present but granted again at a future date but before 1920, which happened in some states) and hence would be in a repeatable events framework where moving in and out of risk is possible. Embedding repeatable events within a competing risks setting would perhaps yield the richest

environment for analysis.

Moving away from suffrage, I could also do an event history analysis of enactment of other kinds of welfare laws, like mother's pensions.

Chapter 4: On Existence of Equilibrium with Endogenous Income

It is well known that in a voting-over-income-taxation game, there exists no pure strategy equilibrium when voters' incomes are exogenous. I prove that this result is true even if the space of tax-schedules is restricted to be marginally progressive only. However in such a setting, I prove that with endogenous income, pure strategy equilibrium exists.

It is interesting to perceive income taxes as equilibrium outcomes of political games played among political parties. The environment is essentially one where political parties or candidates propose income-tax schedules for the whole population with some objective in mind, typically winning the election and collecting some fixed amount of revenue. An individual voter of the electorate then votes for the platform that maximizes his/her own utility. Some rule is employed for determining the winner and his/her announced platform is then implemented (assuming commitment to announcements on the part of the candidates).

It is well known in the literature that there exists no pure strategy equilibrium tax schedule when incomes are exogenous (that is, when voter-utility does not depend on leisure-labor choices so that labor-supply decisions do not respond to tax announcements). Assuming majority rule for determining winners, it can be shown that for any announced tax policy of one party/candidate, it is always possible to construct an alternative one that wins in majority voting over the first one. The main reason for this is that, hurting a smaller fraction of the populace (by imposing a little more tax on them) and favoring a larger fraction

(by lowering taxes for them) and still collecting the required revenue, always seem to be possible in such games (see Klor 2003, for example). The problem is analogous in spirit to the problem of cake sharing among three individuals. It is always possible to take a little bit from

one individual and distribute that among the other two and this allocation gets the support of the two beneficiaries who constitute a majority.

In this chapter, I study the issue of existence of equilibrium schedule with endogenous income (when people do make labor-leisure choices in response to tax announcements). This could be interesting because perhaps, beyond a point, it will not be possible to collect the same revenue anymore by reducing taxes from somebody and taxing others since the high-taxed section can then change their labor choices, and earn the income of the low-taxed section and thus not pay the taxes that were aimed at them. In other words, when people begin to respond by changing labor-supply to tax announcements, it is perhaps not possible to design alternative incentive-compatible tax schedules and still meet revenue requirements beyond a particular distribution of taxes in the population. In that case, that particular tax schedule is likely to be the equilibrium one; no one would be able to do better by announcing some other platform.

I show that with exogenous income, there is no pure strategy equilibrium even when tax schedules are restricted to be marginally progressive only, while with endogenous income there can be pure strategy equilibrium. I in fact prove in this chapter that if the revenue to be collected is high, then there exists an equilibrium and the equilibrium tax schedule is linear (as concave as you can get in the space of convex schedules); if the revenue belonged to some intermediate range of values then also there exists an equilibrium and the equilibrium schedule is strictly convex. For all lower revenue, there exists no equilibrium.

Future Plans:

A more ambitious work will try to relax some of the restrictive assumptions that I have and hence come up with a more general conclusion regarding existence. For example, if I allow for all kinds of tax schedules (both marginally progressive and regressive ones), my conjecture would be that there would still be no pure strategy equilibrium in the endogenous income case as in the exogenous income one. Another relaxation would be using a continuum of voters (I have three income classes, high, middle and low in the chapter).

Chapter 2

Enfranchisement from a Political Perspective

> At the opening of the twenty-first century (and the new millennium), nearly all adult citizens of the United States are legally entitled to vote. What once was a long list of restrictions on the franchise has been whittled down to a small set of constraints. Economic, gender-based, and racial qualifications have been abolished; literacy tests are gone, if not forgotten; residency requirements have been reduced to a matter of weeks; the age of political maturity has been lowered; and the burden of registration has been rendered less onerous. The proportion of adult population enfranchised is far greater than it was at the nation's founding or at the end of the nineteenth century. That there exists a right to vote rather than the privilege of voting is clearly established in law as well as in popular convictions.
>
> Yet getting here has taken a very long time.
>
> -From 'The Right to Vote: The Contested History of Democracy in the United States' by Alexander Keyssar

2.1 Introduction

Democratization, meaning extension of franchise or extension of voting rights, has been a universal phenomenon, down time and across space. The process, though historically witnessed by almost all nations at some point in time, is far from arising from the same kind of incentives every time. The obvious question behind every instance of democratization is that if extension of franchise is against the interests (economic, political and other interests) of the 'elites'[1], then why

[1]Collier [23] refers to three connotations of 'elite': the first is a class concept. It distinguishes the working class from classes "above" it in the social hierarchy. The second is a political concept where 'elite' refers to those having political power which includes not only those participating in the government but also those in opposition. Hence it basically denotes those with voting

would voting rights be extended by them to the 'non-elites'? For example, if democratization implies a poorer median voter and hence greater redistribution, then rich elites might be opposed to extension. They may also be unwilling to share political power with those not belonging to their 'strata'. For example, men might be opposed to sharing political power with women, who were sometimes considered intellectually inferior to men. Hence the question of democratization is interesting and the answer seems to be as diverse as the people undertaking it. Naturally a rich array of models has developed trying to capture these various incentives. Some authors have modeled enfranchisement in response to threats of revolution, war threats, internal split of interests among elites, etc. They also talk about instances of extension that best fit their models of democratization (see the 'Related Literature' subsection below).

In this chapter we model another possible dynamics behind democratization. Our model captures the idea that, even in the complete absence of any threat from the disenfranchised, in the presence of large number of partisan supporters, large rents from office and along with part of the enfranchised seeking their own benefits (which are unrelated to suffrage directly), suffrage can be extended. This model is therefore a model of peaceful extension arising from self-centered interests of the electorate and the parties. Our model fits the experiences of several countries like Chile (1874-91), Sweden (1907-9) etc. Since redistributive repercussions of suffrage extensions are explicitly absent in this model, it can especially explain instances of democratization where redistributive considerations would not have played a big role, like women's suffrage. In fact, women's suffrage in many places, like in the states of USA and Australia, can be explained using this model (see the 'Historical Evidence' section).

---

rights in our context. The third is also a political concept with 'elites' referring to leaders. We primarily use the word in the second sense.

Consider a generic illustration. Let there be two parties, a Liberal Party and a Conservative Party, let only the 'rich' exercise franchise in the beginning, and let there be no other 'external' impetus. Assume, as historically witnessed, that it would be in the political interest of the Liberal party to extend franchise and for the Conservatives not to do so. Now presuming the current electorate (rich elites) is possibly Conservative-aligned, extension becomes a difficult proposition because the Liberals would be less likely to win by proposing extension. But often, franchise has been seen to be extended, not only by the Liberals, but also surprisingly, by the Conservatives as well. Our model proposes a dynamics by which franchise is extended in equilibrium by both the parties. Not only do the Conservatives propose extension in this model, they do so when they are the stronger party in the sense of being more popular. Hence it would seem paradoxical that it proposes extension and moves away from its favorite political agendas, even when it is the more popular party with the current electorate.

The dynamics of the model, which explains this apparent paradox, actually is a story of how the popularity of the Conservative Party proves more of a bane than boon for it. Because it is the more popular party, the current electorate predicts that it will propose policies very close to its own favorites. This may not please a large part of even the current elite voters and they might want extension hoping that this will drive future policies closer to their own. Hence what would have allowed the Conservatives to get as close to their favorite agenda as possible, is precisely what triggers its moving away from it - its popularity with the current voters, which some voters perceive to be detrimental for the policy outcomes. Hence this is a model where extension occurs when the stronger party does not represent the electorate in its policy choices.

We model this idea in the following way. We have a two-period model of election where in the first period the issue to be decided is whether or not to extend

franchise. And in the second period some other policy is under consideration. Both voters and parties have preferences over the second period policy. Moreover, both the voters and the parties are rational and forward-looking and though they do not have any direct preference over who votes, they realize that who gets to vote potentially affects future policies for which they care.

Here we model two effects working in two (possibly opposing) directions for a party when it is considering extension: Firstly, there is a 'direct effect', which means that extending franchise (possibly benefiting some voters and hurting others) affects the probability of winning in this period. Secondly, there is an 'indirect effect', which means that franchise extension today will potentially affect future policies for which the parties care.

Ordinarily, with diverse party preferences, the 'indirect effect' is likely to be positive for the party whose next-period favorite becomes more achievable due to extension in period 1 (the Liberals). It is likely to be negative if extension drives outcomes away from one's favorite. The 'direct effect', again can go either way - some voters today would like extension because it drives tomorrow's outcome closer to their favorite while the opposite happens for others. A party proposes extension if the positive effect outweighs the negative so that the sum of the effects is positive. In our model, under large rents from office and one of the parties enjoying an advantage with respect to distribution of partisanships, we get large and positive 'direct effects' for both the parties, so that both the parties propose extension (even when the indirect effect is negative). Hence the model explains franchise extension, by not only Liberals but also Conservatives, who presumably would have to move away from their favorite political agenda (negative indirect effect) due to extension.

The positivity of the 'direct effect' can be explained as follows: Under certain conditions (which includes evenly balanced partisan preferences, and large rents from office, among others) there is convergence of policies in the second period away from the median voter's policy. So possibly a majority of voters today, would want to support extension, if that means coming closer to their own favorites. In other words, extension means future policies coming very close to their own, and hence they support extension. The other voters who were happy before, obviously do not support extension, but on the whole the parties find that extension increases their probability of winning this period (positive 'direct effect'), when the total marginal gain in utility outweighs the total marginal loss in the current electorate.

Since there is no explicit use of taxes and transfers in my model, it can also be applied to cases of extension of suffrage, that was not likely to have large redistributive repercussions, like women's suffrage. For example, consider the enfranchisement of the women in America (see Scott and Scott [78] for a vibrant description of the suffrage movement). The facts, in a nutshell are these - suffrage for all women was federally mandated in 1920[2]; but states had power to grant suffrage to women before that. About half of the states did grant women's suffrage prior to 1920, in between 1890 to 1920, and it is this phenomenon that we are going to look at more closely. It turns out that there was a definite geographical pattern to the "suffrage" states - they were all the western states or Mountain states (Wyoming, Colorado etc). Now if one looks at the "power distribution" or "party-presence" at that time, one finds that power seemed to be most closely balanced between the two leading parties (Republicans and Democrats) in precisely the Western states. Loosely speaking this was because, after the Civil War (which ended in 1865), both northerners and southerners started moving west

[2]World War I and women participation in substituting the male workforce at home is often thought to be the plausible reason for this, see Ticchi and Vindigni [84].

(which was largely uninhabited before). The northerners were mostly Republican supporters while the southerners were mostly Democrats and the medley settlers turned out to be pretty much evenly balanced (in the sense of large number of partisan supporters in favor of each party). Moreover, possibly a part of the existing electorate also wanted extending of franchise (like some ideas at the time suggesting men wanting to enfranchise women to attract them in the tardily female-populated and dominantly young-male-populated mining states of the west, see the 'Historical Evidence' section). Also with large federal funding going into these states, rents from office were likely to be very high. So as this example is suggestive of, our model tries to capture the idea that strong partisan bases, coupled with interest of part of the electorate, might prove conducive to the enfranchisement of the disenfranchised[3] (see chapter 3 of this dissertation for a related empirical work on partisan competition and women's suffrage in the United States).

Related Literature:

A seminal body of research talks about enfranchisement due to threats from within like revolution (see Acemoglu and Robinson [1], [2], [3], and Conley and Temimi [24]). In [3] for example, Acemoglu and Robinson develop a model that explains democratization due to threat of revolution. In [1] again, they develop a dynamic theory which explains how democracies are created and why some democracies consolidate (like Britain) while some go on and off democracy (like Argentina) as a result of revolutions etc, which again result from inequalities in the economy. So the more the inequality in income distribution, the more the chances of revolution and the more unconsolidated the democracy. Conley,

[3]See, Keysser [49] and Munger [66] for arguments that the western states were "more competitive" than the rest of the country at that time. The North, Northeastern were more Republican-strongholds while the South seemed more of a Democratic stronghold.

Temimi, also argue in a similar vein, that when the preferences of the enfranchised and the disenfranchised group conflict, and the latter pose a threat (revolution, unrest, civil disobedience), then suffrage is granted by the former to the latter when the threat is considerable. However these theories are less applicable to the kind of examples we have in mind since there is no evidence of any potentially threatening political upheaval in those cases.

Ticchi and Vindigni [84] argue, that to induce the masses to fight harder in a war, political rights are extended as a commitment to favorable redistribution in future. However, this theory cannot explain women's and men's suffrage in situations without threats of war. Llavador and Oxoby [56], models enfranchisement arising out of elite split. When the elite (enfranchised group) is split and so is the working class (the disenfranchised group), the faction that comes to power tries to extend franchise to those aligned ideologically with itself (like the liberals granting voting rights to the industrial workers only, while the conservatives, with interests more aligned with those of landlords and peasants, seeking universal suffrage since the latter were at the lower end of the income distribution and voting rights were income contingent at that time). However that does not conform to reality in many instances. For example, if men or women (the split in our case) had been more aligned with one or the other party (the Republicans or the Democrats), we should have seen extension of franchise to women in either the northern or the southern American states where each of the parties had strongholds which obviously did not happen. In fact, it happened in those states where the parties had the closest partisan competition and franchise was extended to women under both Republican and Democratic regimes in these western states[4].

[4]Himmelfarb [44] was also of the opinion that partisan competition between the Conservatives and Liberals led to franchise extension in Britain. It was like "parties competing against each other in a miserable auction with the constitution being 'knocked down to the lowest bidder'. Again it was held that Disraeli (Conservative) "was partial to the residuum (poor) because it was most likely to be managed and *exploite* by the Conservative party... But if this

Lizzeri and Persico [55] models enfranchisement arising out of elites wanting to steer away politicians from serving the electorate with targeted redistribution, towards making investments in public policies with diffused benefits, especially with an exogenous increase in the value of public goods (like sanitation, health facilities etc, in the face of rising epidemics like cholera). The model is used to explain increased public spending during Britain's "Age of Reform". Our model is similar to [55] in that franchise extension, is to some extent, elite-driven. However, the interest of the elites in our model can be very general. And there are no specific redistributional and investment-related forces acting on the voters and the parties in our game (see Aidt [4] and [5], for limited and conditional evidence of redistribution following suffrage extension). Hence our model helps to explain women's suffrage in many cases, which are generally agreed to have no serious redistributive repercussions. Other theories of democratization include those of Ghosal and Proto ([37]), Jack and Lagunoff ([47]), and Bertocchi ([10]) among others.

There exists a nonformal political science literature about political competition leading to democratization (see Himmelfarb [44] for argument in the British case and Collier [23] for a general argument)[5]. The theory (similar to Llavador and Oxoby's in[56]) predicts that a party might extend franchise with the expectation that the newly enfranchised will return the favor by voting for their party in return. In this model, parties do not extend franchise with the hope or expectation of getting the votes of the new voters. In fact, with strong partisan divisions in the present electorate, they are pessimistic or at best skeptical regarding the party

---

were the only consideration, the Conservatives would surely have tried to extend the suffrage to the residuum in the counties, where the opportunities for management and exploitation were even greater; yet there was not even the most wistful hint of such a proposal."

[5]Political sociology also points towards a similar spirit (see Lipset [54]): "a stable democracy requires the manifestation of conflict or cleavage so that there will be struggle over ruling positions, challenges to parties in power, and shifts of parties in office."

inclinations of groups who had not voted earlier[6]. Especially for women, since most of them were expected to be close to some men personally and the latter were equally divided between the parties, the politicians would have apprehended that the new voters would vote quite similarly to the existing ones[7]. This theory also has often not been true historically with the party extending franchise often ending up losing in the next election[8]. And rationally speaking, once enfranchised, the new voters are likely to vote for the party proposing more favorable platform for the policy at hand, not the party that had enfranchised them. In our model extension takes place to gain votes from part the *current* electorate, not a future one. Hence here extension takes place in the midst of high uncertainty as to the outcome of extension. This was certainly true in case of women's suffrage in the Western American states, for example.

The rest of the chapter is organized as follows: Section 2 lays down the model in general terms. Section 3 covers equilibrium analysis of period 2 election. Section 4 covers equilibrium analysis of period 1 election while the interpretation of conditions are carried on in Section 5. Section 6 elucidates the main ideas using some specific functional forms for utilities, distributions, etc. Section 7 gives historical examples of franchise extension to which this theory seems to apply. Section 8

[6]The confusion is clear prior to women's suffrage in New Zealand. Atkinson [9] writes "... the question (was), how would the women vote?... Ballance (Liberal) privately feared that most of the 'forward' and 'connected' women - those more likely to enroll and vote- would back his conservative opponents, resulting in a 'crushing and overwhelming defeat for the Liberal Party'." Again, Grimshaw [41] writes about the Conservative sentiment: "...women would never vote conservative. 'They will be the first to ask the reason why a few men should be absolutely rolling in wealth, and wasting more in a day than would keep some of these families for a month'." Again for other conservatives " women was undoubtedly a conservative element in the community, the upholder of established institutions of Church and State..."

[7]This actually turned out to be true historically. In New Zealand for example, contrary to their fears, the Liberals won immediately after enfranchising the women and "... there is no evidence that women's votes had any impact on the overall result" (Atkinson [9]). In the case of America, Flexner [36] says: "women have shown the same tendency to divide along orthodox party lines as male voters."

[8]Conservatives lost the election in Britain immediately after extending the franchise.

contains the conclusion.

2.2 The General Model

This is a two period model with elections in each period. In the period 1 election, the issue to be decided is who has the right to vote in the period 2 election. In the period 2 election, some other policy is to be chosen. It is convenient to start by describing what happens in period 2 since this part is fairly standard.

2.2.1 Period 2:

Let $p \in [0, 1]$ be the policy to be decided in this period.

Preferences of the Parties:

There are two parties A and B. Each party cares about two things: policy and rents from office. Winning in any period allows a party to exploit rents from office, say R. Let their 'utility' from policy (p) be given by welfare functions $W_A(p)$ and $W_B(p)$ respectively (where, for existence purposes we need the maximizers of W_A and W_B to be sufficiently dispersed, see Grossman and Helpman [42], and Persson and Tabellini [69], for proof). Also without loss of generality let us assume, as we will in the example in Section 2.6 later, that $W_A(.)$ is maximized at a higher p than $W_B(.)$. Let the parties propose p_A and p_B respectively, in this period.

Preferences of the Voters:

Voters belong to a population which is represented by the $[0, 1]$ interval (but the electorate may or may not be the whole of the population, as described in the next subsection). Heterogeneity of the voters can be thought of as representing difference in incomes with voter 1 being the richest and voter 0, the poorest. In

the context of gender, voters can be thought of as differing in education levels (or physical strength) with voter 0 being the most illiterate (or physically least strong) and voter 1, the most educated (or the strongest), so that we can think of having women in the lower end of the voter spectrum[9].

Voters care about two things: policy and ideology. The ideology can bias them towards one party or another. (As we will see in the next section, bias is modeled to be probabilistic and drawn from a distribution.) More rigorously, we have the following: Let j be a generic voter. j'th voter's total utility from A winning in period 2 is given by $v_{jA2} + u_j(p_A)$ where v_{jA2} can be variously interpreted as j's partisan or ideological preference for party A or alternatively his liking for the fixed policies of party A in period 2, while $u_j(p)$ is his utility from the pliable policy, which is p this period. u_j is assumed to have a unique maximizer (preference is assumed to be single-peaked over policy p) and for simplicity we assume that the maximizers are increasing with j. Hence for any two voters j and k,

$$j > k \Leftrightarrow \arg\max u_j(p) > \arg\max u_K(p).$$

For simplicity, we could assume, as we do in Section 2.6, that $\arg\max u_j(p) = j$. This just means that smaller j's like smaller p's while higher j's like more p. Similarly, j's total utility from B winning in period 2 will be $v_{jB2} + u_j(p_B)$.

Now if policy space is interpreted as tax rates then we have the poorest voter wanting a low tax rate for themselves while the richest could want a higher tax rate for the not-so-well-off. Alternatively, if the policy space was to represent 'level of alcohol consumption', then 'higher' voters (the men), might prefer higher levels of p while 'lower' voters (the women) might prefer lower values of p.

[9]Due to various social norms, women were denied formal education for a long time.

We now describe the 'non-standard' part of our model.

2.2.2 Period 1:

In period 1, only part of the population is enfranchised. However, they may decide to extend the franchise before policy decision is made in the next period. We model it as indirect democracy: the two parties each propose enfranchisement decision; then the voters vote on the one they prefer. After that, there is a new election (period 2 election) where policy p is decided (again in an indirect democracy).

More rigorously, in period 1, the electorate is a fraction $[\bar{m}, 1]$, $\bar{m} < 1$, of the whole population. Call it the 'original' electorate. This can variously be interpreted like people voting above some income threshold (the richer population above $\bar{m}$), or only 'men' voting.

In the first period, the issue to be decided is whether or not to extend franchise (call this issue m). Hence each party proposes either extension or no extension in period 1. Formally, parties propose $m_i \in [0, \bar{m}]$, $i = A, B$. For example, parties can choose to lower income threshold or give suffrage to women. Existing voters vote and the winner is selected. (Notice, the lower is m, the greater is the extension, so that $m = 0$ means full extension of suffrage.) The winning party i implements its proposed policy platform, so that the period 2 electorate is $[m^*, 1]$ where $m^* = m_i$.

Hence suppose the franchise is extended to $m^* \leq \bar{m}$. That means, all types above m^* are entitled to vote in the next period. Call $[m^*, 1]$ the 'extended' electorate. The extended electorate will determine policy p in the second period election while it is the 'original' electorate which determines if the franchise should be extended or not.

Preferences of Parties and Voters:

We assume that no one (voters or parties) has any inherent preference over this first period policy, m. That is, no one is directly concerned about who gets to vote in the next period[10]. However, as described, everyone has a preference over the second period policy, p. So notice that voters and parties will have an induced preference over m, since who votes in the second period affects equilibrium p policy over which they have preference[11].

2.2.3 Timeline:

The timeline is as follows:

Period 1 election -
1. Parties propose $m_i \in [0, \bar{m}], i = A, B.$
2. Original electorate votes.
3. Winner is selected.

Period 2 election -
4. Parties propose $p_i \in [0, 1], i = A, B.$
5. Extended electorate votes.
6. Winner is selected.

2.3 Subgame Equilibrium

We begin by analyzing the game backwards, from the period 2 election.

---

[10]Historically, this seems to be the case in American states, for example. McDonagh and Price (see [60]) writes "... neither party was formally involved in woman suffrage until 1916...".

[11]This also seems to be the case historically. In America, for a more recent example, "The Liquor Dealers were convinced that women, if they could vote, would bring about prohibition" (see Scott and Scott [78]). Hence their opposition to women's suffrage. For a more classical example, p can be thought of as tax-rate and hence peoples' preference over it.

2.3.1 Period 2 Election:

Given preferences of the voters in subsection 2.1, j votes for A if A's winning gives j higher total utility than B's winning i.e. if

$$v_{jA2} + u_j(p_A) \geq v_{jB2} + u_j(p_B). \tag{2.1}$$

Let $v_{jA2} - v_{jB2}$ be defined as v_{j2}. Hence, rearranging terms, j votes for A if

$$v_{j2} \geq u_j(p_B) - u_j(p_A). \tag{2.2}$$

Now let us assume, as in probabilistic voting models, that actual values of v_{j2} are not known to the parties but they have a common prior belief about the distribution of v_{j2}. Moreover, this distribution of v_{j2} depends on another aggregate shock in the economy, which is assumed to be random. Parties again have a common prior belief of the distribution of this aggregate shock. Due to this random aggregate shock affecting individual partisan distributions, the vote share of each party is random, so that each party wins with some probability. Let $P_{A,2}(p_A, p_B, m)$ be the probability that party A wins in period 2 when it proposes p_A and the opponent proposes p_B, and the electorate is m ($\leq \bar{m}$, the winner of last period's election). B wins with the remaining probability, $P_{B,2}(p_A, p_B, m)$. We can get an explicit expression for $P_{A,2}(p_A, p_B, m)$ once we make specific assumptions about the distribution of v_{j2}.

Given the preferences of the parties, and their probabilities of winning, we can define the "expected utilities" of the parties in this period as follows:

$$U_{A,2} = P_{A,2}(p_A, p_B, m)(R + W_A(p_A)) + P_{B,2}(p_A, p_B, m)W_A(p_B) \tag{2.3}$$

$$U_{B,2} = P_{B,2}(p_A, p_B, m)(R + W_B(p_B)) + P_{A,2}(p_A, p_B, m)W_B(p_A) \tag{2.4}$$

where, if a party wins, his total utility is its welfare from implementing its own proposed policy, plus some positive rents from office-holding, R. In case he loses,

which happens with the remaining probability, he gets no rents from office and since his opponent's policy is implemented, gets welfare from p evaluated at the opponent's proposed value. Now party A maximizes (2.3) by choosing p_A while B maximizes (2.4) by choosing p_B. Assume that all the conditions that guarantee existence of an interior unique optimum solution, hold. Then the first order conditions for maximization will be as follows:

$$\frac{\partial P_{A,2}(p_A, p_B, m)}{\partial p_A}(R + W_A(p_A) - W_A(p_B)) + P_{A,2}(p_A, p_B, m)W'_A(p_A) = 0 \quad (2.5)$$

$$\frac{\partial P_{B,2}(p_A, p_B, m)}{\partial p_B}(R + W_B(p_B) - W_B(p_A)) + P_{B,2}(p_A, p_B, m)W'_B(p_B) = 0. \quad (2.6)$$

Now, as usual, the FOCs have natural interpretations. Consider (2.5). Suppose A considers changing p_A slightly. Then this has two effects. Firstly, for every voter j, this might change whether j votes for A or B which in turn changes the vote share for A and hence $P_{A,2}(p_A, p_B, m)$ changes. For example, we might think that when p_A falls, for many of the voters with high 'j', $u_j(p_A)$ will fall so that $P_{A,2}(p_A, p_B, m)$ might be adversely affected, while relatively low 'j' voters might like it, helping the probability of winning. Suppose, for expositional ease, the net effect is a positive change in the probability. Now with every unit gain in the probability of winning, the utility gained is $(R + W_A(p_A))$ while the utility achievable even if A didn't win is $W_A(p_B)$. Hence the net total utility that A is concerned with, when probability of winning is changing by one unit, is $(R + W_A(p_A) - W_A(p_B))$. Hence the total gain in net utility due to the probability of winning changing by $\frac{\partial P_{A,2}(p_A,p_B,m)}{\partial p_A}$, is given by $\frac{\partial P_{A,2}(p_A,p_B,m)}{\partial p_A}(R + W_A(p_A) - W_A(p_B))$.

Secondly, however, if A does win, its welfare from implemented policy p_A is lower by $W'_A(p_A)$ (since we've assumed that A likes a higher p) and this happens with probability $P_{A,2}(p_A, p_B, m)$. Hence expected "loss" from lowering p_A slightly will be $P_{A,2}(p_A, p_B, m)W'_A(p_A)$. Hence as always, A lowers p_A as long as the 'positive' or 'benefit' part outweighs the 'negative' or 'loss' part and stops when they are exactly equal. Similarly we can interpret B's FOC.

Let the resulting optimal policies that solve (2.5) and (2.6) be $p_A^*(.,m)$ and $p_B^*(.,m)$ where the '.' in the argument of the policies denote the distributional parameters of the model and we explicitly mention 'm' since it is actually endogenous in our whole model, though a parameter in the second stage election game. Plugging them back in (2.3) and (2.4), we can write $U_{A,2}(m)$ and $U_{B,2}(m)$ as the equilibrium expected utilities of the parties in period 2 when the electorate is m. This was a standard election model. Let us now turn to the first stage.

2.4 Equilibrium in Period 1 Election

$[\bar{m}, 1]$ people vote in period 1. Let parties propose (m_A, m_B) in period 1. As we said before, current voters don't directly have any preference over who votes in period 2 but since who votes in period 2 affects the equilibrium outcomes in period 2 (p_i^*'s, $i = A, B$) over which voters have preferences, they have induced preferences in period 1 over policy m. We model this preference as follows. Voter j votes for that party whose winning is better for him considering both, his partisan preference for the parties in this period and his expected total utility from period 2 (the latter comprising of utility from the p policy and his partisan preference in period 2).

Now let $\mathcal{B}_2$ be the set of all possible aggregate shocks in period 2 in the economy and let a generic element of $\mathcal{B}_2$ be b_2. Let the set of aggregate shocks in the economy such that A wins in period 2, given $[m, 1]$ people are voting (m is the winner of period 1 election), be denoted by

$$\mathcal{B}_2(A, m) := \{b_2 : \text{A wins in period 2}\}.$$

Hence $\mathcal{B}_2(B, m) := \mathcal{B}_2 - \mathcal{B}_2(A, m)$. Now if A wins in period 1, then $[m_A, 1]$ people vote in period 2. Hence today (without discounting) the expected utility of voter

j from A winning now is

$$v_{jA1}+\int_{\mathcal{B}_2(A,m_A)}[u_j(p_A^*(.,m_A))+E(v_{jA2}|b_2)]d\mu(b_2)$$
$$+\int_{\mathcal{B}_2(B,m_A)}[u_j(p_B^*(.,m_A))+E(v_{jB2}|b_2)]d\mu(b_2).$$

But $\int_{\mathcal{B}_2(A,m_A)} d\mu(b_2) = P_{A,2}(.,m_A)$. Hence j votes for A if total utility from A winning today exceeds that of B winning today. i.e.

$$\begin{aligned}
&v_{jA1}+P_{A,2}(p_A^*,p_B^*,m_A)u_j(p_A^*(.,m_A))+P_{B,2}(p_A^*,p_B^*,m_A)u_j(p_B^*(.,m_A))\\
&+E(v_{jA2}|\mathcal{B}_2(A,m_A))+E(v_{jB2}|\mathcal{B}_2(B,m_A))\\
\geq&v_{jB1}+P_{A,2}(p_A^*,p_B^*,m_B)u_j(p_A^*(.,m_B))+P_{B,2}(p_A^*,p_B^*,m_B)u_j(p_B^*(.,m_B))\\
&+E(v_{jA2}|\mathcal{B}_2(A,m_B))+E(v_{jB2}|\mathcal{B}_2(B,m_B)).
\end{aligned}$$

Letting $v_{jA1}-v_{jB1}$ to be v_{j1}, rearranging terms, simplifying and defining

$$\begin{aligned}
&[P_{A,2}(p_A,p_B,m_B)u_j(p_A^*(.,m_B))+P_{B,2}(p_A,p_B,m_B)u_j(p_B^*(.,m_B))\\
&+E(v_{j2}|\mathcal{B}_2(A,m_B))]\\
-&[P_{A,2}(p_A,p_B,m_A)u_j(p_A^*(.,m_A))+P_{B,2}(p_A,p_B,m_A)u_j(p_B^*(.,m_A))\\
&+E(v_{j2}|\mathcal{B}_2(A,m_A))]
\end{aligned}$$

to be $= L_j$, we get, j votes for A if $v_{j1} \geq L_j$. Again, the distribution of v_{j1} is assumed to be affected by another aggregate economy-wide random shock in period 1 (drawn from $\mathcal{B}_1$ with generic element b_1, say, both defined analogously like $\mathcal{B}_2$ and b_2 respectively) that makes the vote share of the parties random. Hence let $P_{A,1}(m_A,m_B)$ be the probability that A wins in period 1, when A proposes m_A and B proposes m_B. Now, like the voters, parties realize that franchise today determines the policy tomorrow over which they have preferences. Hence A and B maximize the following "expected utilities" in this period by choosing m_A and m_B respectively:

$$U_{A,1}=P_{A,1}(m_A,m_B)(R+U_{A,2}(m_A))+P_{B,1}(m_A,m_B)U_{A,2}(m_B) \tag{2.7}$$

$$U_{B,1}=P_{B,1}(m_A,m_B)(R+U_{B,2}(m_B))+P_{A,1}(m_A,m_B)U_{B,2}(m_A). \tag{2.8}$$

Here expected utility of A = Pr(A wins today)x(what A gets if he wins today) + Pr(A doesn't win today)x(what A gets if he doesn't win today).

If A wins today, he not only gets rents from office today but also expected utility tomorrow evaluated at m_A, given A has won today. Similarly, if A loses today, he gets no rents from office today and tomorrow's expected utility evaluated at m_B, given B has won today. Plugging $P_{A,1} = 1 - P_{B,1}$, the FOCs will be:

$$\frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A}(R + U_{A,2}(m_A) - U_{A,2}(m_B)) + P_{A,1}(m_A, m_B)\frac{\partial U_{A,2}(m_A)}{\partial m_A} \quad (2.9)$$

$$\frac{\partial P_{B,1}(m_A, m_B)}{\partial m_B}(R + U_{B,2}(m_B) - U_{B,2}(m_A)) + P_{B,1}(m_A, m_B)\frac{\partial U_{B,2}(m_A)}{\partial m_B}. \quad (2.10)$$

The solution will either be at a corner or an interior: parties will extend franchise (i.e. lower the cutoff m) as long as (2.9) and (2.10) are negative at $\bar{m}$; there will be full extension (up to 0) if (2.9) and (2.10) are negative at 0; will not extend if they are positive at $\bar{m}$ and we will have an interior optimum if (2.9) and (2.10) can be solved when set equal to 0.

2.5 Intuitive Interpretation

Like in period 2, we can invoke an intuitive interpretation of the above conditions. We can see a 'direct effect' (the first term in each of the above expressions which is itself product of two terms) and an 'indirect effect' (the second term in the above expressions, also a product of two terms) of changing an announcement slightly.

Direct Effect:

This captures the effect of extension of franchise on the probability of winning today and the associated gain (or loss) in utility. Notice that, since extension of franchise is represented by lowering m, a negative sign of the derivative $\left(\frac{\partial P_{.,1}(.)}{\partial m}\right)$

will actually imply a positive 'direct' effect, in the sense that extending franchise (lowering m) increases the probability of winning today.

Take party A, for example. By decreasing his announcement of m_A slightly (extending franchise), he directly affects his probability of winning this period (how, is explained later in this section). Suppose, for expositional ease, the probability of A winning this period increases as A lowers m_A. Now with every unit increase in the probability of winning, the utility gained is $(R + U_{A,2}(m_A))$ while the (expected) utility A would be getting anyway, even if he didn't win this period is $(0 + U_{A,2}(m_B))$ (corresponding to B winning this period). Hence the net total utility that A is concerned with, when probability of winning is changing by one unit is $(R + U_{A,2}(m_A) - U_{A,2}(m_B))$. Hence the total gain in net utility due to the probability of winning changing (increasing) by $\frac{\partial P_{A,2}(m_A,m_B)}{\partial m_A}$, is given by $\frac{\partial P_{A,2}(m_A,m_B)}{\partial m_A}(R + U_{A,2}(m_A) - U_{A,2}(m_B))$.

Indirect Effect:

This captures the effect of extension of franchise on the party's utility tomorrow due to a change in tomorrow's platforms.

If A does win in this period, his expected utility in period 2 changes by $\frac{\partial U_{A,2}(m_A)}{\partial m_A}$, as he changes m_A slightly now. However he wins in this period by $P_{A,1}(m_A, m_B)$. So his expected change in next period's expected utility if his announcement of franchise changes today slightly is $P_{A,1}(m_A, m_B)\frac{\partial U_{A,2}(m_A)}{\partial m_A}$.

In fact, we would be interested in seeing when will the 'direct effect' help extension.

Signing the Direct Effect

(Heuristics behind positive 'direct effect')

1. Under certain conditions, the second period policies converge above the median voter's policy. Intuitively, as rents from office gets large, and with large number of partisan supporters in favor of each party, both the parties tend to propose 'interior' policies, away from their favorites (A's policy is pulled down while that of B is pulled up). But now suppose A is the relatively stronger party (in the sense that it is likely to have partisan distributions biased in its favor), then it does not have to move too much away from its favorite. That is, B's policy is pulled up more than A's is pulled down, so that both are above the median's.

2. Since the policies are further above the median, possibly a majority of the current voters (the 'lower' voters) wants a lower policy. An extension of the electorate does just that - it results in lower equilibrium policies. A's policy will still be above B's, and both will be above the new median's but the new median will be lower, and the resultant policies will be much closer to the ('lower') current electorate's favorites. Hence the relatively 'low' voters find it in their interest to try and shift the electorate so that the resulting policies tomorrow are very close to their own favorites.

In short, as A lowers m_A, the change in utility might be better for some voters (the 'lower' voters, here), and worse for others. The total effect is better from the point of view of the parties when the total marginal gains in utility of the 'lower' voters is bigger than the total marginal losses in utility of the 'higher' voters due to a smaller equilibrium policy tomorrow. Hence it means that a party's probability of winning by extending franchise increases (positive direct effect), as long as it can do it without hurting current voters too much on the whole. Similarly, B's vote share and probability of winning are likely to increase

(see appendix .1 Democratization-A for a proof of vote share and probability of winning, moving in the same direction) when A's do[12] (since voters vote for any party proposing extension).

Hence we can resolve the apparently paradoxical result of extension of franchise by a Conservative party which is 'strong'. This apparent 'strength' precisely turns out to be its biggest hindrance in getting its favorite policy. A large part of the voters forsee that the Conservative party will announce policy too close to its own favorite and further from theirs so that they want extension to get closer to their own favorites.

Regarding the term multiplied to the change in probability in the 'direct effect' term, $(R + U_{A,2}(m_A) - U_{A,2}(m_B))$, when the difference in expected second-period utilities is bounded and R is large, the sign is likely to be positive. Hence the 'direct effect' is likely to be positive as long as the probability of winning in period 1 rises as franchise is extended (cutoff lowered).

Signing the 'Indirect Effect':

Notice that for party A, it is likely to be negative with A's own favorite being a 'high' p so that extending franchise and winning (that's likely to drive down p), is likely to make A worse off. However this effect should be positive for party B whose favorite p is 'low'.

Hence franchise is extended (m is extended downwards) as long as the sum of the direct and the indirect effects is positive when evaluated at $\bar{m}$. Notice that once the direct effect is positive, both the effects are likely to be positive for party B,

[12]Notice that if the parties were just maximizing their probabilities of winning, they would make the same optimal choice and we would have, what is called, 'policy convergence', even with probabilistic voting, in the absence of any intrinsic party preferences over m.

while for A, the indirect effect is likely to be negative. Hence whether A extends franchise or not depends on the relative strengths of the two effects. It extends when the positive direct effect outweighs the negative indirect one. In order to characterize conditions under which we can have large positive direct effects and therefore 'extension', we turn to a more concrete example.

2.6 An Example

In line with the specification laid down in the general model above, let the j'th voter's preference over p, $\forall j$, be:

$$u_j(p) = -(p - p(j))^2$$

where $p(j)$ is the most-preferred policy of j and utility falls the further one goes from $p(j)$. We also assume that the function $p(j)$ is differentiable with $p'(.) > 0$ which means that the more rightward a voter, the higher the p he prefers. For simplicity and without loss of generality, let $p(j) = j$, $\forall j$. Hence the utility from p for the j'th voter is:

$$u_j(p) = -(p - j)^2.$$

Let the parties welfare from policy p be represented by the following:

$$W_A(p) = p$$

$$W_B(p) = 1 - p$$

which means that A's favorite policy is $p = 1$ while that of B is $p = 0$. Let also voting be probabilistic in both the periods. We begin by analyzing the game backwards, from the period 2 election.

2.6.1 Period 2 Election:

According to earlier notation, j votes for A if $v_{j2} \geq u_j(p_B) - u_j(p_A)$. Now let us assume, as in probabilistic voting models, that v_{j2} follows some distribution

which is known to both the parties. Again, for simplicity, we follow the standard distributional assumption that $v_{j2}|b_2 \sim U\left[\frac{b_2}{f} - \frac{1}{2f}, \frac{b_2}{f} + \frac{1}{2f}\right]$, $\forall\ j$. This gives a uniform distribution with height f which measures the diversity of preferences for the fixed positions of the parties in the voting population and a shift parameter b_2 which measures popularity of the party's fixed positions on the whole, and which is assumed to be random. Hence a lower f would mean a 'flatter' distribution with possibly many partisans on either side and smaller number of 'neutrals' (if the distribution is more or less around 0, depending on b_2). b_2 measures the overall popularity of the fixed positions of the parties in this period. For example, if the pliable positions of the parties are same, then j votes for A if $v_{j2} \geq 0$. Now if $b_2 > 0$ the whole distribution is shifted towards being positive so that A is more popular (since $v_{j2} > 0$ is more likely to be realized) whereas if b_2 is negative, the distribution is shifted towards being negative which means B is more popular, based on fixed positions only.

Thus the probability that j votes for A, given announcements (p_A, p_B), is given by

$$P(v_{j2} \geq u_j(p_B) - u_j(p_A)) = 1 - F_{v_2}(p_A^2 - p_B^2 - 2j(p_A - p_B))$$

(after substituting for the utility functions to obtain the last line), where $F_{v_2}(.)$ is the cumulative distribution function of the above uniform distribution (assumed same for all j). Let for any pair of announcement (p_A, p_B), A's vote share[13] in

[13]In more general mathematical notation, suppose for every individual j the decision function for voting for A looks like follows:

$$(X_j|b_2)(\omega) = \begin{cases} 1 & \text{if } j \text{ votes for A,} \\ 0 & \text{if } j \text{ votes for B.} \end{cases}$$

So a natural definition of the actual vote share for A should be

$$V_{A,2}(p_A, p_B, m)(\omega) = \frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj.$$

And a natural definition of the expected vote share for A (the object of interest for us) should

period 2, given m people are voting, be given by $V_{A,2}(p_A, p_B, m)$, where

$$V_{A,2}(p_A, p_B, m) = \frac{1}{1-m}\int_m^1 \{1 - F_{v_2}(p_A^2 - p_B^2 - 2j(p_A - p_B)\}dj.$$

Now under certain conditions (which the conditions of proposition 1 imply), we

---

be:

$$V_{A,2}(p_A, p_B, m) = E\Big(\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj\Big).$$

However, what I really have as the expression for $V_{A,2}(p_A, p_B, m)$ is

$$V_{A,2}(p_A, p_B, m)(\omega) = \frac{1}{1-m}\int_m^1 E(X_j|b_2)(\omega)dj.$$

(since I integrate over all j, the probability that j votes for A, given b_2). Notice that going from

$$E(\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj) \quad \text{to} \quad \frac{1}{1-m}\int_m^1 E(X_j|b_2)(\omega)dj$$

is not problematic (assuming we make the right convergence assumptions) but that from actual vote share, to expected vote share, is. i.e. the following is problematic:

$$\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj \quad \text{to} \quad E(\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj).$$

To see why, suppose the voting population is discrete ($j = 1, 2, ..., n$). Then, by the strong law of large numbers we could say that

$$\frac{1}{n}\sum_{j=1}^n X_j(\omega) \to E(\frac{1}{n}\sum_{j=1}^n X_j(\omega))$$

(almost surely). (And by a change in the order of summation we could have obtained $E(\sum_{j=1}^n X_j(\omega)) = \sum_{j=1}^n E(X_j(\omega))$). A continuous analogue should have been the following:

$$\int_m^1 (X_j|b_2)(\omega)dj \to E(\int_m^1 (X_j|b_2)(\omega)dj)$$

(almost surely). (And then by change of order of integration we could have obtained what I have used as expected vote share i.e. $\int_m^1 E(X_j|b_2)(\omega)dj$. However, such a law of large numbers is not available in the literature. For one, $\int_m^1 (X_j|b_2)(\omega)dj$ may not even be a measurable function so that expectation will not even be defined. We prefer to bypass this problem by defining $V_{A,2}(p_A, p_B, m)$ as the expected vote share, i.e. $E(\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj)$ and not trying to derive it from the actual vote share $\frac{1}{1-m}\int_m^1 (X_j|b_2)(\omega)dj$. Note that we assume the usual convergence theorems to make the change in the order of integration meaningful and actually work with the quantity $\frac{1}{1-m}\int_m^1 E(X_j|b_2)(\omega)dj)$ as the expected vote share.

So to summarize, the mathematical difficulty is this: First, the natural definition of actual vote share, $\int_m^1 (X_j|b_2)(\omega)dj$, may not be well-defined, so that the notion of expected vote share is not meaningful. Second, even if it is, it is difficult to get to a quantity like $E(\int_m^1 (X_j|b_2)(\omega)dj)$ from it (since an analogue of the law of large numbers is absent). Again, both these problems don't arise with finitely many voters.

get

$$V_{A,2}(p_A, p_B, m) = \frac{1}{2} + b_2 - f(p_A^2 - p_B^2 - (1+m)(p_A - p_B)).$$

And the vote share of B in period 2 when m people are voting, $V_{B,2}(p_A, p_B, m)$, is naturally $1 - V_{A,2}(p_A, p_B, m)$. Now since b_2 is random, the vote share of each party is random. Let $b_2 \sim U[-\gamma + \delta, \gamma + \delta]$. Then, by our notation in the general model $\mathcal{B}_2 := U[-\gamma + \delta, \gamma + \delta]$. (Hereafter, we will abuse notation slightly and not differentiate between the random variable and its realizations.) Notice that $\delta = 0$ would mean that the "fixed-position-popularity" parameter is equally likely to be in favor of either party whereas a negative or a positive δ would mean that the resulting distribution would be biased in favor of one party over another. Again, by our notation in the general model:

$$\mathcal{B}_2(A, m) = \{b_2 : V_{A,2}(p_A, p_B, m) > \frac{1}{2}\}.$$

And hence using the distributional assumption of b_2, probability that A wins in period 2 is given by:

$$P_{A,2}(p_A, p_B, m) = \frac{1}{2} + \frac{\delta}{2\gamma} - \frac{f}{2\gamma}(p_A^2 - p_B^2 - (1+m)(p_A - p_B))$$

(again assuming $F_{b_2}(t) = \frac{t+\gamma-\delta}{2\gamma}$ holds and the condition for this is again implied by the conditions in proposition 1). Again, B wins with the remaining probability. As defined in the general model, parties maximize "expected utilities" by choosing p_A and p_B respectively. Here the timing of this stage is as follows:

There might be a second way around the problem, which is formulating the model a little differently. Let all voters have an identical bias v_2 (no j subscript, because the bias is common to all voters), which is uniformly distributed on an interval $[\frac{\delta-1}{f}, \frac{\delta+1}{f}]$. Hence once v_2 is realized, the voters between $[0, 1]$ are again ordered (but with a shift) and the probability that a party wins is the probability that the aggregate bias realization (v_2) is such that the median voter votes for that party (since all the others on one side vote similarly) and it wins. Working out this way, we get all the results as above (with gamma equal to 1). For example, the probability of winning of party A in period 2 turns out to be

$$P_{A,2}(p_A, p_B, m) = \frac{1}{2} + \frac{\delta}{2} - \frac{f}{2}(p_A^2 - p_B^2 - (1+m)(p_A - p_B)).$$

And of course, the problem of law of large numbers does not arise.

1. The parties announce (p_A, p_B). Neither v_{j2} nor b_2 is known at this point.

2. Both v_{j2} and b_2 are realized and all uncertainty (individual and party levels) is resolved. Only b_2 needs to be realized for aggregate uncertainty to be resolved.

3. Elections are held and winner is selected.

Hence the political equilibrium of this stage of the game can be summarized as follows:

Proposition 1 *Define* $\Delta := R^2 + \frac{4\gamma}{f}$. *Let* $f > 0$, $\gamma > 0$, *and* $f < \frac{1}{2}\Big(-\gamma + \delta + \frac{1}{2}\Big)$, *and* $\left(\frac{1+m}{2} + \frac{R}{4}\right)^2 \Delta > \left(\frac{1}{4}\Delta - \frac{\delta}{2f}\right)^2$ *hold. Then* $(p_A^*(.), p_B^*(.))$ *is a Nash equilibrium of the period* 2 *election where*

$$p_A^*(., m) = \frac{1+m}{2} - \frac{R}{4} + \frac{\delta}{2f}\frac{1}{\sqrt{\Delta}} + \frac{1}{4}\sqrt{\Delta}$$
$$p_B^*(., m) = \frac{1+m}{2} + \frac{R}{4} + \frac{\delta}{2f}\frac{1}{\sqrt{\Delta}} - \frac{1}{4}\sqrt{\Delta}.$$

(Please see appendix .2 Democratization-B for proof.)

(Since m is a parameter at this stage of the game, the equilibrium choices of p and consequently $U_{A,2}$ and $U_{B,2}$ are functions of m and the other parameters of the model, namely f, γ, δ and R.)

Observation 1: If we take first order approximations of the terms under square roots and let $R \to \infty$ and $f \to 0$ such that Rf is finite, we get the following:

$$p_A^*(., m) = \frac{1+m}{2} + \frac{\delta + \gamma}{2Rf}$$
$$p_B^*(., m) = \frac{1+m}{2} + \frac{\delta - \gamma}{2Rf}.$$

Now if δ is assumed to be greater than γ, we get both the policies to be greater than the median $\frac{1+m}{2}$ (where median is $m + \frac{1-m}{2} = \frac{1+m}{2}$), with p_A^* bigger than p_B^*.

Observation 2: Note that the conditions in the above proposition implies $\delta > -\frac{1}{2}$ which is only what we need for an interior solution to exist. But what

we are going to need for extension of franchise in the first period would be somewhat stronger which is $\delta > \gamma$. As far as distribution of partisan preferences are concerned, notice that $\delta > \gamma$ implies that the aggregate shock is always positive which means that the individual partisan distributions of v_{j2} will always be biased towards being favorable for A. Thus this condition implies an ex-ante advantage of party A (the Conservative party in our set up).

Observation 3: Notice that f has to be small in some sense (that is, a lot of partisans on either side), for an interior solution to exist.

2.6.2 Period 1 election:

$[\bar{m}, 1]$ people vote in period 1. And let $v_{j1}|b_1 \sim U\left[\frac{b_1}{f} - \frac{1}{2f}, \frac{b_1}{f} + \frac{1}{2f}\right]$, $\forall\ j$. Notice that the distributions of $v_{j1}|b_1$ and $v_{j2}|b_2$ are the same $\forall\ j$. According to notation introduced earlier, $V_{A,1}(m_A, m_B)$, the vote share of party A in period 1, is given by:

$$V_{A,1}(m_A, m_B) = \frac{1}{2} + b_1 - \frac{f}{2\gamma(1-\bar{m})}\int_{\bar{m}}^{1} L_j dj.$$

Like in period 2, let $b_1 \sim U[-\gamma + \delta, \gamma + \delta]$. Then, by our notation in the general model $\mathcal{B}_1 := U[-\gamma + \delta, \gamma + \delta]$ (so notice that the economy-wide random shock is drawn from the same distribution in both the periods). Using the distributional assumption for b_1, $P_{A,1}(m_A, m_B)$, the probability that 1 wins in period 1, will be

$$P_{A,1}(m_A, m_B) = \frac{1}{2} + \frac{\delta}{2\gamma} - \frac{f}{2\gamma(1-\bar{m})}\int_{\bar{m}}^{1} L_j dj.$$

Now, in FOCs for period 1 we need to sign terms like $\frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A}$. We can show that $\frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A}$ will be positive, negative or zero, depending on the sign of $\frac{\partial V_{A,1}(m_A, m_B)}{\partial m_A}$ (see appendix .1 Democratization-A for proof) where

$$\frac{\partial V_{A,1}(m_A, m_B)}{\partial m_A} = -\frac{f}{1-\bar{m}}\int_{\bar{m}}^{1} \frac{\partial L_j}{\partial m_A} dj.$$

Hence the sign of $\frac{\partial V_{A,1}(m_A,m_B)}{\partial m_A}$ depends on the sign of $\int_{\bar{m}}^{1} \frac{\partial L_j}{\partial m_A}$. In case an interior optimum exists, we will have an m_A^* that makes $\frac{\partial V_{A,1}(m_A,m_B)}{\partial m_A} = 0$. Here, since we are interested in 'extension', we will find conditions under which the derivative of the objective function (the FOC) is non-positive when evaluated at $\bar{m}$ (as m falls, 'utility' rises). In particular, for the derivative of the vote share to be negative (which also means that the derivative of the probability of winning is negative which in turn would mean that the derivative of the objective function is negative), we need $\int_{\bar{m}}^{1} \frac{\partial L_j}{\partial m_A} dj$ to be non-negative. The following proposition summarizes the conditions.

Proposition 2 *Suppose Proposition 1 holds. Suppose $R \to \infty$ and $f \to 0$ such that Rf is finite. Then "extension" up to at least m is a Nash equilibrium of the first period election when*

$$\gamma + Rf(\bar{m} - m) < \delta.$$

Furthermore, $m_A^* = m_B^* = 0$ or "full extension" is a Nash equilibrium if, in addition,

$$\bar{m} < \min\Big\{\frac{\delta - \gamma}{Rf}, 1 - \frac{\delta + \gamma}{Rf}\Big\}.$$

(Please see appendix .4 Democratization-D for proof.
Also see appendix .3 Democratization-C for a proof that there exists a Nash Equilibrium.)

Observation 4: Note that the full-extension condition in the above proposition implies that $\bar{m} < \frac{1}{2}$ since,

$$\bar{m} < \frac{\delta - \gamma}{Rf} < \frac{\delta + \gamma}{Rf} < 1 - \bar{m}.$$

Hence it means, that for full-extension, the size of the current electorate should be pretty large (at least half of the whole population).

The above propositions, tell us that full extension is a Nash equilibrium of this political game in the following scenario:

When f is small (tends to 0) i.e. partisan preferences are more towards a particular party (and less neutral); when δ is greater than γ, which means that party A is more favorably placed than party B as far as fixed positions of the parties stand[14]; when gains from office get large; and when the current electorate is large, then extension is proposed by both parties.

Here one would think A to be more reluctant towards extension since its own favorite p policy is 1, and a bigger electorate means a smaller value of p_A^* (see the expression for p_A^* where m enters positively, so that larger electorate means smaller m and hence smaller p_A^*). Moreover the existing electorate is also the "upper" part of the population whose overall tendency would be to vote favorably towards higher p, in line with what A wants. However, in this model, A's apparent advantage (the partisan bias in its favor) turns out to be its biggest foe. Since A has a partisan advantage, the electorate knows that A's p policy will be very 'high'. Not only that, the electorate knows that with winning very important for both the parties (R being high), B's policy will also be 'high' (lower than A's, but above the median) i.e. the two policies cannot be very far apart. Hence a majority of the electorate supports extension since that surely pulls down both the platforms, and they get policies closer to their own favorites.

Graphically, we get the following: The second period policies converge (asymptotically) to policies greater than the median. A's policy converges to $p_A^*(\bar{m})$, as in the diagram, and B's policy converges to $p_B^*(\bar{m})$, lower than A's, but both above the median. So notice that more than half the electorate would like a lower

[14]Regarding extension of franchise in Britain, Himmelfarb [44] says "There is no doubt that party advantage played a large part in the strategy of both Conservatives and Liberals."

p policy, i.e. gets better off as the policies come closer to $\frac{1+\bar{m}}{2}$. Also notice that how further the policies are from the median depends on the bias-parameters, δ and γ.

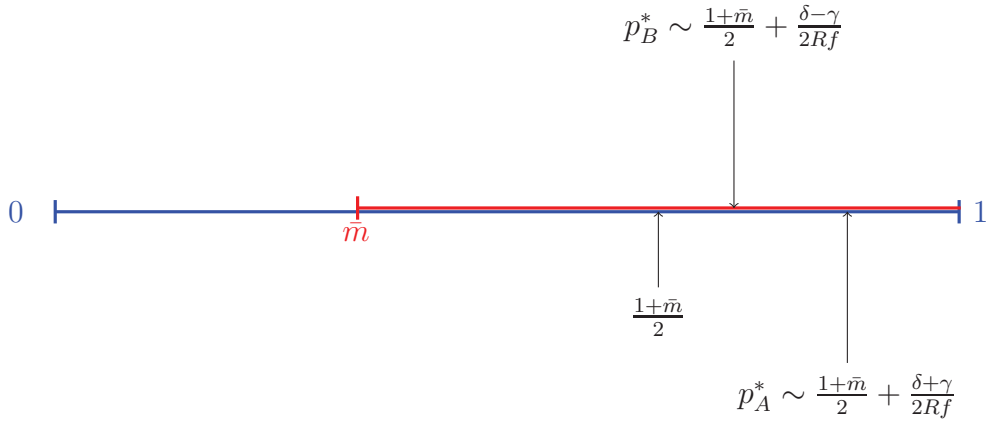


Figure 2.1 Before extension

Now suppose the electorate in the second period is $[m, 1]$. This drives both the possible period-2 policies, p_A^* and p_B^*, down, possibly below $\frac{1+\bar{m}}{2}$, and definitely closer to a majority's preferred policy. Hence extension increases probability of winning in the first period for any party proposing extension.

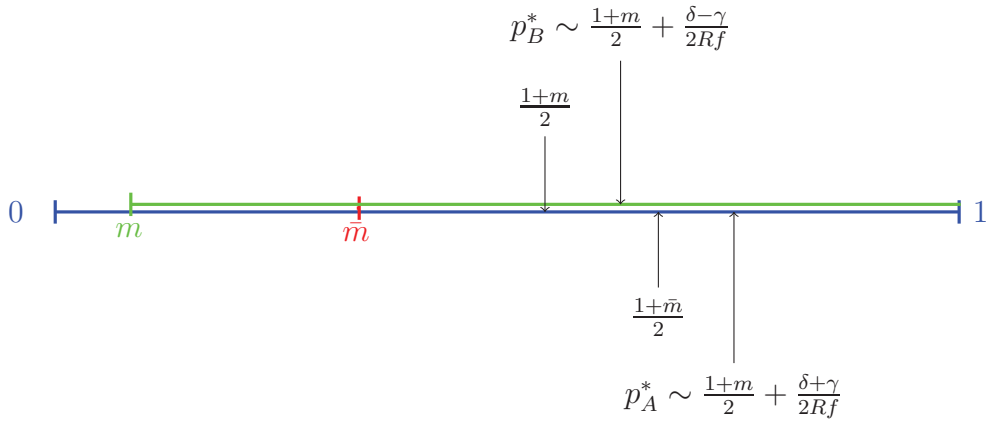


Figure 2.2 After extension

Also notice that the extent of extension depends on how further from median the policies without extension would be, which in turn, depend on the bias-parameters, δ and γ.

2.7 Historical Evidence

We would like to highlight the main features of our model in the following examples. To recall, the main features include: initially limited suffrage, strong partisan competition between parties (in the sense of large number of partisans in favor of each party), suffrage extension initiatives of both Liberals and more importantly Conservatives, and part of the enfranchised supporting suffrage extension. Moreover, we give possible examples of direct and indirect effects, wherever we could find them.

Sweden (1907-1909):

Background: The Parliament Act of 1866 abolished the system of the four Estates of Nobles, Clergy, Burghers and Farmers and established the Swedish Parliament that was divided into a First Chamber and a Second Chamber. Both the Houses were to have equal competence and powers in all matters. Abolition of the Clergy Estate reduced the authority of the Lutheran Church in political affairs but there was hardly any change in the type of members elected to the Parliament. It was almost as if the two upper Estates had been joined together to form the Upper House and similarly for the Lower one. There was little change in the composition of the Houses for the next forty years.

Who were enfranchised: Five percent of the 4 million inhabitants (20% of adult males) in 1870 possessed vote to the Lower House.

Who supported suffrage extensions:

The growth of democratic sentiments among the population reflected the changes that occurred since 1866. With a full-fledged industrial revolution, people began to demand a change in the oligarchic structure that had replaced the estates. Booming mining and metal industries increased the population of industrial workers. The Universal Suffrage Association was formed. The idea of universal suffrage attracted two small but expanding groups, the liberals and the socialists

whose goal was to achieve *Folkriksdag* or People's Parliament. Hence, the social effects of the industrial revolution manifested themselves in the growth of a new middle-class interested in suffrage, temperance and Free Trade. Collier [23] sees democratization reforms in Sweden, for example, as being partly driven by the political 'ins' (those that already have franchise, like part of the electorate in our model).

> ...[A]t the time of the reform, the working class was already at least partly enfranchised, with the result that a labor-based party had already won seats in the parliament... (Collier [23])

Partisan Competition and Conservative Initiative: Partisan competition between the Liberals and the Conservatives in Sweden was also significant. Rustow [74] says:

> The protracted fight (between the Liberals and the Conservatives) stirred the passions to an unprecedented degree, and... both the times the conservatives yielded at the eleventh hour. Earlier they had sacrificed the estates so as to preserve oligarchy. Now they endorsed manhood suffrage for the chamber (1907) to safeguard their position in the senate,...

Assuming the divide in the legislature is representative of the partisan divide in the electorate, there is clear evidence of strong partisan support in favor of each party in the electorate. In 1912, for example, the distribution of seats in the Lower House among the Conservatives and Liberals was 64 and 102 respectively, while that in the Upper House was 86 and 52 respectively, so that overall the Conservatives had 150 seats while the Liberals had 154, with the Conservatives stronger in the Upper House and Liberals in the Lower.

Karl Staaff and his Liberal government proposed manhood suffrage in 1906 but it was not passed in the Second Chamber. It was finally passed by a Conservative government in 1907/9, led by Arvid Lindman. He not only accepted Staaf's program for a Second Chamber franchise for men over twenty-four but also offered

a limitation of the multiple votes in local elections to a maximum of forty per person. For the first time, a Conservative had attacked the plutocratic basis of the senate, which even the Liberals had hitherto left untouched, and the bill went far beyond even what the Left expected. The lifting of income qualifications for lower house voters at a single stroke doubled the electorate. This represents what has been often called "a great compromise"(Rustow [74]).

Direct and Indirect Effects: Rustow [74] writes:

> In the great compromise of 1907 they (the Conservatives) were able to force adoption of proportional representation, which appreciably slowed the parliamentary ascent of the Liberal and the Socialist parties... (This) led to a consolidation of conservative forces...

Hence initially the Conservatives seemed to have made an electoral gain (positive direct effect). However, enfranchisement of thousands of lower class persons weakened Conservative strength in the Parliament and they lost their virtual monopoly in the First Chamber. In the first local election, the combined strength of the United Rightists and the Moderates dropped from 133 to 86 in the First Chamber.

Chile (1874-1891):

Background: Before 1874, income and property qualifications restricted the franchise. But the 1874 reforms resulted in a threefold increase in suffrage from 1872 to 1878 (Valenzuela [87]). It basically removed all property qualifications, though literacy requirements were retained.

Partisan Competition and Conservative Initiative: Conservatives and Liberals appear to be quite evenly balanced around the period that democratization took place in Chile. In the Chilean Chamber, the number of Conservatives were 25 while the number of Liberals were 22. In the Municipal elections, the percentage of votes in favor of Conservatives was about 26% while that in favor of Liberals was about 20.4%.

Arturo Valenzuela ([87], [30]) writes that conflict among the Chilean elites revolved primarily around ideological questions such as the role of the church and control over government resources. Conservative landed elites were highly discontented with government expanding its jurisdiction. Again, the Radical party representing anticlerical and mining interests also pressed for suffrage and democratization. Hence the Conservative party, opposed to central control and seeking further liberalization of suffrage found an unlikely ally in the liberals over the suffrage issue.

> [T]he Right sought to advance its interests through the democratic electoral process... From a position of strength in Congress, the conservatives, together with Radicals and ideological liberals [successfully pressed for a series of reforms]. (Valenzuela [87])

Collier [23] also writes about the growing power and more liberal orientation of "career civil servants" who became more inclined towards liberal policies, desiring increased autonomy within the state. This corresponds to part of the electorate in our model, in whose favor extension of franchise would work. The liberals proposed the 1874 reform which was pushed in the legislature by the conservatives.

Direct and Indirect Effect: Valenzuela [87] says, "while the Conservatives initially gained from electoral reform and were able to dominate the politics of the Parliamentary Republic (positive direct effect), they could not foresee that the electoral reform would soon benefit a new group of parties with far different agendas (negative indirect effect)". In the Chilean context, this happened through nitrate production (mining, not a conservative product) giving rise to a host of other ancillary industries, spawning a new working class leading to the emergence of the Communist party in Chile, opposed to the conservatives.

Italy (1912):

Background: After the Italian Risorgimento (unification or 'resurgence') between 1859 and 1870's, franchise was restricted on the basis of property and literacy. In 1871, 72.96% of the population was illiterate and after income restriction only 1.98% of the population was enfranchised. The Reform Act of 1882 relaxed only the property restriction. However since most of the peasantry and majority of artisans and workmen, were still illiterate, the percentage of enfranchised rose from 2.18% to 6.97% of the population.

In 1912, a new Reform Act extended franchise to all men who had served in the armed forces and all who were above thirty. No literacy qualification was required. This was almost universal suffrage.

Partisan Competition and Conservative Initiative:

Many constitutional groups and factions joined to form an amorphous Conservative block. The various cliques that called themselves after their leaders - Depretian, Crispinian, Rudinian, Zanardellian, Giolittian, and so on. Universal suffrage (1912) was a Conservative initiative under Giolitti's 'great ministry' of 1911-1914 and such a political maneuver might well have been brought about by the rise of the 'Nationalists' between 1910 and 1914. Nationalists, standing for long-lasting interests of the country appealed to many like the professional middle class and students. They wanted social justice and redemption of the proletariat at home through redistribution of world's wealth acquired at wars. Hence they combined Left and Right ideologies, exactly what the Giolitti ministry granted.

The Conservative forces also seem to be equal in strength to the Liberal forces in Italy. Around the period of democratization, the Conservative party got about 19.1% of the votes in the Italian Chamber of Deputies election, while the Liberal Parties got about 19.4% votes.

Direct and Indirect Effects: Universal suffrage that came into effect in 1913 kept the old single-member constituencies but widened the male suffrage to include most of the nations proletariat and peasantry. This helped to put a brake on forces of radical change (benefit, positive direct effect) like the extreme Left.

However the costs of colonial war far outran all expectations and showed no signs of lessening which put an immediate end to welfare programs at home, which became suicidal in the face of universal manhood suffrage (negative indirect effect). After the first election with broadened suffrage in 1913, the premier's career was in the decline.

Women's Suffrage: To the best of our knowledge, no serious modeling efforts have been made to capture circumstances of extension of voting rights to women. In [2], page 18, Acemoglu and Robinson, in fact write,

> ...when the roles began to change as women entered the workforce, women also obtained voting rights... the mechanisms that we propose better describe the creation of male suffrage than the extension of voting rights to women.

Again according to the *'enlightenment'* theory, 'enlightenment' made the elites realize it was not fair and just for a large fraction of the population to be deprived of representation. Hence the above authors write in [2],

> ... the enfranchisement of women in Britain in 1919 and 1928 appears mostly due to changes in the society's approach to women.

However 'change in gender-roles' or 'enlightenment' theories seem to predict a smooth incorporation of women's suffrage into political lives of people at that time which is far from the truth. There was considerable effort made by suffrage organizations to wrench the right of franchise out of unwilling hands. Moreover men wanted to extend voting rights to women when women were believed to be allies in their own objectives (like prohibition, anti-child labor laws for the

progressives in USA), perhaps not out of enlightenment. We look at the following cases closely:

In States of the United States:

Partisan Competition: During the post Civil War period, the Rocky Mountain West was characterized by intense two-party competition and power seemed to be closely balanced. The reason for this balance seemed to be resulting from and reflecting the preferential divide of the electorate itself. These regions were sparsely populated at the time of the Civil War and most settlers trickled in during subsequent decades - among the migrants were the Southerners, many of them ex-Confederates, and perhaps their preferences leaned towards the Democrats; while those moving into these states from the Union base were likely to be Republican supporters. Women's suffrage was also granted first in these states (see chapter 3 of this dissertation for a related empirical work on partisan competition and women's suffrage in the United States). The following figures make the point most clearly. Notice that the dark-colored states in the maps are the ones with more partisan competition and early enactment of suffrage.

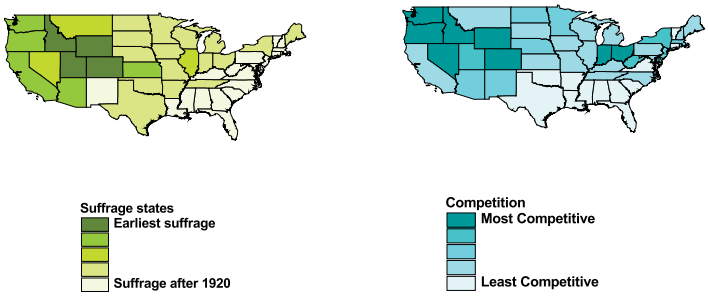


Figure 2.3 Suffrage states and political competition in the United States

We look at Idaho more closely.

Idaho: Idaho became the third state in the USA to grant women's suffrage in 1896. Idaho appears to be a 'two-party system' according to Ranney and Kendall [70] (and others) where, in most of the elections

(a) two parties have shared the bulk of the votes and public offices between them,

(b) the winning party has gained a majority of the votes and offices, and

(c) the two dominant parties have alternated in winning majorities.

Conservative Initiative: Like in our model, both the Liberals (Democrats) and Conservatives (Republicans) wanted to extend franchise to women. The following planks were adopted by the state conventions of Idaho:

The Democratic convention: "We recommend to the favorable consideration of the voters of the state the proposed constitutional amendment granting equal suffrage, believing that this great question should receive the earnest attention of every person as an important factor in the future welfare of the state."

The Republican convention: "We favor the amendment to the constitution of the state proposed by the late Republican legislature, including equal suffrage for men and women, and recommend their adoption."

In New Zealand: New Zealand women were given the right to vote in Parliamentary elections in 1893. Hence it became the first national state in the world to allow women to vote.

Who wanted women enfranchisement: The Women's Christian Temperance Union, and several Suffrage Leagues supported suffrage. Many men, together with the institutions they managed, also lent support - the most potent being the politicians themselves. Ready sympathy came from those nonconformist churches where elements of women's equal rights had already been accepted. The Congregational Church, the Methodists, the Baptists, and Salvation Army members gave backing to the cause. There was support from the Presbyterian Church and the Church of England as well.

Conservative Initiative: At the time when women's suffrage was granted in New Zealand, the Government was a Liberal one while the Opposition seemed

to be more Conservative. However both the groups seemed to favor extension of suffrage to women. The majority of the Opposition members held woman to be undoubtedly a conservative element in the community, the upholder of established institutions of Church and State, as the following newspaper clipping hints:

> We hold very strongly that the effect of the admission of women to the suffrage will be distinctly Conservative, and the only possible corrective to the evils attendant on universal manhood suffrage. *(The Evening Post)*

Once the suffrage legislation passed, "both the Liberal government and the opposition subsequently claimed credit for the enfranchisement of women, and sought women's newly acquired votes on these grounds" (Atkinson [9] 84-94).

In Australian States: Regarding the political situation in South Australia, the first state to grant women's suffrage in 1894, Audrey Oldfield says (see [67]),

> ...politics were largely factional, with Governments forming and reforming in response to internal alliances. This created instability - there were twenty-eight Governments in the first twenty years...

Hence there may not have been 'inter-party' partisan competition but some form of political competition in the guise of factional competition definitely seems to have been present. We look at the Australian state of Western Australia more closely.

Western Australia (1899):

Background: In 1887 the British Colonial Office agreed to draft a Constitution that gave a fully elected Legislative Assembly. For the Assembly, there were property qualifications for electors and elected, with plural voting and no payment of members. With only 5900 males eligible to vote, the first legislators were drawn from the wealthy and influential. John Forrest became Premier and was to remain so till 1900.

Conservative Initiative: In 1893, there was growing unrest about the Assembly franchise among those who were streaming into Western Australia from other

more democratic Australasian colonies in response to the great gold discoveries. Conservatives like J. Cookworthy, R. Scholl, wanted amendments that proposed widowed, single and divorced women with the same property qualifications like men, to have a vote. They seemed to want this to offset the democratization of the Assembly by increasing the propertied vote in the Council. Thus Western Australia granted (propertied) women's suffrage in 1899, thereby becoming the second Australian state to do so (after South Australia). writes the following regarding women's suffrage in Western Australia:

> ...passage of female suffrage legislation by the conservative government of John Forrest in an attempt to strengthen the urban conservative vote against the emergent labour movement of the goldfields. (Magarey, see Daley and Nolan [29])

2.8 Conclusion

Franchise extension has been a nearly omnipresent phenomenon in the political history of the world. Here we develop a model that explains franchise extension solely through the presence of elite-pressure, large rents from office and evenly-balanced partisan competition. The model is especially helpful in explaining cases of democratization that are not accompanied by widespread redistributional repercussions. In case of women's suffrage, in most cases it has been gradual with the male and the richest being the first to vote. Much later were voting rights also extended to women. Moreover the magnitude of the suffrage movements were less than that of 'revolutions' so that the 'revolution' theories cannot be applied to women's suffrage movements in general. Though many attempts have been made to model the 'male' version of franchise extension, unfortunately, we don't find any systematic efforts to model extension of franchise to women. This model can fill in that gap.

A natural avenue for future research would be to explore extension of franchise

in other countries and see whether this model can be extended to explain those situations. Some Islamic countries, for example, seem to fit the 'partisan competition' aspect of our model. Saudi Arabia, for instance, which has an absolute monarchical form of Government has no women's suffrage while Bahrain, which is a constitutional monarchy implying a much liberal political ambience, in the presence of different political parties, has women's suffrage. An empirical research in this direction should also be rewarding. Also historical evidence suggests that interest groups (industries like breweries etc) often tried to influence politicians on issues of suffrage etc. It would be nice to incorporate such lobby influences on extensions.

Chapter 3

Partisan Competition and Women's Suffrage in the United States

3.1 Introduction

> "Good morning, sister. You taught us and trained us in the way we should go. You gave us money from your hard earnings, and helped us to get a start in the world. You are interested infinitely more in good government and understand politics a thousand times better than we, but it is election day and we leave you at home with the idiots and Indians, incapables, paupers, lunatics, criminals and the other women that the authorities in this nation do not deem it proper to trust with the ballot; while we lordly men, march to the polls and express our opinioins in a way that counts."
>
> - From Belle Kearney, A Slaveholder's Daughter (St. Louis: The St. Louis Christian Advocate Press, 1900), pp 111-112

Women in many of the states of the USA did not enjoy the same voting rights as men until 1920 when the Nineteenth Amendment to the Constitution granted women's suffrage. Passed by Congress on June 4, 1919, and ratified on August 18, 1920, the nineteenth amendment granted women the right to vote with the following stipulation:

> The right of citizens of the United States to vote shall not be denied or abridged by the United States or by any State on account of sex.
>
> Congress shall have power to enforce this article by appropriate legislation.

Some thirty years previously, however, Wyoming entered the Union with a constitution granting women full voting rights (1890). Several other states also gave women the vote before 1920. A timeline of granting full suffrage to women is

given in Table 3.1[1].

State	Year
Wyoming	1890
Colorado	1893
Idaho	1896
Utah	1896
Washington	1910
California	1911
Oregon	1912
Kansas	1912
Arizona	1912
Montana	1914
Nevada	1914
Nebraska	1917
New York	1917
Michigan	1918
South Dakota	1918
Oklahoma	1918

Table 3.1: Timeline of full suffrage for women

The first states to grant suffrage were the Mountain states of Wyoming, Colorado, Idaho and Utah followed by the Pacific states (Washington, Oregon, California).

[1]Some of these states had granted suffrage earlier while they were territories. For example, Wyoming granted women's suffrage as a union territory in 1869 and it was still in place with statehood in 1890. Utah had women's suffrage as a territory since 1870 which was annulled by Congress in 1887, but by referendum was put back in the constitution when Utah was admitted to statehood in 1896. Washington had woman suffrage twice by the enactment of territorial legislature but lost it by court decisions. Also, Arizona's suffrage enactment in 1910 came into effect with statehood in 1912. And Montana granted women's suffrage way back in 1887 as a territory. So the table lists *states* where full suffrage was granted to women (and not repealed) before 1920.

In this paper we want to study systematically the causes that might have led these states to enact women suffrage earlier than others, prior to 1920.

Many theories have been put forward to explain early suffrage in the western states. One of the popularly held beliefs regarding early suffrage in the western states is the adverse female-male sex ratio in those states. During the 1880's the west was largely populated by young males and the idea was that suffrage rights would make those states sufficiently attractive for young women to come and settle in. On the other hand, Keyssar [49] attributes suffrage to "the unusual political circumstances" that prevailed in the handful of states (Wyoming, Colorado, Idaho, and Utah) where suffrage was achieved. He talks about the strength of the People's Party being crucial to the 1893 success of women's suffrage in Colorado. In Utah, also, he says that the enfranchisement of women was probably linked to politics of the Mormon territory with a tradition of polygamy (see Keyssar [49], for example). Strength of suffrage organizations and suffrage movements were also believed to be important in bringing about suffrage rights early to the west (see McCammon and Campbell, [59], for example).

Other scholars have held that demographic factors can explain early western suffrage. Western states, for instance, had few blacks so that extending suffrage to women did not raise the racial issues that it did in other states, like demands for enfranchisement of the blacks following women's suffrage (see Keyssar [49], for example). Still others have argued that changing of gender roles to be the most crucial factor in bringing about suffrage (see McCammon and Campbell [59], for example). The idea in this case was that in the western states with harsh frontier conditions of life, women were required to participate on an equal footing with men in many activities and hence the gradual degeneration of traditional gender roles paved the way for suffrage laws.

Despite the large literature on women's suffrage, the quantitative analysis to date

has been inconclusive (see Miller [61], for example). This paper attempts to fill in this gap by hypothesizing and testing certain factors that might have led to suffrage enactment. I propose a new hypothesis that partisan competition between the Republicans and the Democrats in the western states contributed significantly to early suffrage in those states. These western states were characterized by 'close' two-party competition (in the sense that each party had many partisan supporters) which led me to hypothesize that partisan competition, appropriately captured, might explain part of the early suffrage in these states. More precisely, my hypothesis (based on the related chapter on democratization of this dissertation) is that states where Democrats and Republicans were very closely balanced were the ones most likely to enact suffrage while the states in which one of the parties was a stronghold were least likely to enact suffrage.

The intuition is that in the face of stiff partisan competition, and part of the electorate wanting extension (like young males trying to attract women to the tardily female-populated states of the west, or progressives trying to find supporters in women for their causes like prohibition and anti-child labor legislation), one or both the parties extend franchise to please part of the electorate wanting extension thereby increasing their chances of winning.

Hence I test my hypothesis along with some of the ones previously put forward. I do so within an event history analysis framework. I find that partisan competition indeed played an important role in early enfranchisement of women in the west. I also find that some of the popularly believed factors like adverse female-male sex-ratio, and higher percentage of urban population (possibly capturing the percentage of educated professional and progressive middle classes) contribute significantly and positively towards suffrage enactment (as required by the theory). I also find significant 'diffusion' of suffrage enactment in the western states which seems plausible given the geographic clustering of the states in one particular

region. There is also evidence of significant positive duration dependence.

3.1.1 Related Literature:

This work builds on and contributes to both the theoretical and the empirical literature on suffrage extension.

Along the first strand, there is a huge body of literature that looks at suffrage extension in general, but relatively few that look at the extension of suffrage to women in particular. One theory models enfranchisement as a result of threats imposed by the disenfranchised group, especially when the interests of the two groups conflict a lot (see Acemoglu and Robinson [1], [2], [3], and Conley and Temimi [24]). The threats could be in the form of civil disobedience or popular unrest. The theory predicts that if the threats are too costly (implying that the disenfranchised group is very strong and it is more costly for the elites to curb the upheaval than to grant the agitators political rights), franchise is likely to be granted. This might well be true in some contexts but the problem in applying this to the variation in the timing of women's suffrage in American states is that there is not much evidence in favor of widely varying strengths of the suffragists across states that could result in some states granting suffrage but not others.

For example, an organized suffrage movement was almost absent in Wyoming, the first state to enact full suffrage for women. There was no suffrage in Connecticut where the first women's suffrage organization was established prior to the 19th amendment. There was almost equal suffrage organization membership in the west and the south (which are regions respectively where suffrage laws seem to be most and least successful). The movement mobilized early in the eastern states but enactment did not follow. King, Cornwall and Dahlin [51] find that the strength of the suffrage movement increased the likelihood of bill introduction but not bill passage (see Miller [61] and for some of these ideas).

A second category visualizes enfranchisement due to threats from outside like war, which need mass mobilization of armies (see Ticchi and Vindigni [84]). Ticchi and Vindigni's argument is that with the threat of war, mass armies are required; to induce them to fight harder to prevent a loss, their stakes from loss have to be raised through monetary redistribution; but this cannot always be done credibly and so political rights are granted as a commitment to favorable future income distribution. Again, this theory cannot explain granting of voting rights to women only in the western states since there was no threat of war only for those states. However World War I and women participation in substituting the male workforce at home is often thought to be the plausible reason for universal women's suffrage in 1920.

Bertocchi [10] has a theory specifically addressing the extension of franchise to women and also tests the theory empirically. She looks at enfranchisement of women under the assumptions that men are richer than women, women display a higher preference for public goods and women's disenfranchisement carries a societal cost. She proxies for the cost of disenfranchisement with the presence of Catholicism, which is associated with a more traditional view of women's role and thus a lower cost. Higher preference for public goods is proxied by the availability of divorce, which implies marital instability and a more vulnerable economic position for women, leading to higher demands for Government provisions. The gender wage gap is proxied by the level of per capita income. The model predicts that women's suffrage is positively affected by per capita income and negatively by the presence of Catholicism and the availability of divorce. These predictions are then empirically tested. First of all, the assumptions are very restrictive and probably not true only for the western states and not others. Moreover I include partisan competition explicitly which is not considered in this model.

The results of other empirical investigations pertaining to early suffrage in the

western states have been most inconclusive (see Miller [61]). The only robust finding seems to be the positive effect of percentage of women working in non-agricultural activities (see King, Cornwall and Dahlin [51]). The closest to my work is perhaps that of McCammon and Campbell [59] who study the same question of early suffrage in the western states, using event history analysis, and find that mobilization of suffrage movements in the west, and 'political' and 'gendered' opportunities in those states, play the most important roles for suffrage enactment. Here they capture 'political' opportunities by dummies for endorsements from Republican/Democratic parties or Third parties, and categorical variables for procedural ease[2]. To capture 'gendered' opportunities, they include proportion of lawyers and physicians who were women, percentage of all college and university students who were women, and the number of women's organization.

From the point of view of methodology, they differ (which might explain differences in some of the results) in that they run separate regressions with different sets of variables (like one with only political variables, one with only demographic variables and so on) and then run the final regression with the variables that turn out to be significant in the separate (sub-) regressions. However, while running separate regressions, variables that are being left out are not really being controlled for, so that significance of a political variable, for example, when only two

[2]This matters for them because their data starts in 1866 when many states were actually territories, with much less complex reform procedures. However only 3 of the 12 states that were territories, during the suffrage years passed women's suffrage during their territorial years (Utah, Washington and Wyoming) while the others did so as states, indicating that procedural ease may not have been that vital, to begin with. Even then this issue is avoided here by starting the window in 1880 and following full suffrage enactments in *states only.* Again, there was procedural differences within states but little evidence that this was important for laws. For example, Delaware did not require referendum vote on women's suffrage (all other states did), yet did not have full suffrage. Similarly, Michigan required only one legislative vote prior to a referendum, but had suffrage much later in 1918. In any case, I believe that unless the force of political competition was playing a role in the background, procedural simplicity alone would not let reforms pass, and similarly, with strong competition, procedural difficulty would be little hindrance to passing laws.

other political variables are considered might well not be there if other demographic variables (of another regression) are included. Hence in my analysis, all relevant variables are included for all specifications.

The rest of the paper is organized as follows: the next section presents my partisan competition hypothesis and the resultant empirical predictions. Section 3 presents other plausible hypotheses. Section 4 introduces the empirical methods. Section 5 discusses the data. Section 6 presents analysis and results. Section 7 checks for robustness using alternative measures of competition. Section 8 concludes.

3.2 Partisan Competition and Suffrage Extension

> A stable democracy requires the manifestation of conflict or cleavage so that there will be struggle over ruling positions, challenges to parties in power, and shifts of parties in office.
>
> - 'The Political Man: The Social Bases of Politics' (1960) by Seymour Martin Lipset

In the post Civil war period most districts in the West were closely balanced between the parties. The Rocky Mountain West is characterized by intense two-party partisan competition and power seems to be closely balanced. The reason for this balance seems to be resulting from and reflecting the preferential divide of the electorate itself. These regions were sparsely populated at the time of the Civil War and most settlers trickled in subsequent decades – among the migrants were the Southerners, many of them ex-Confederates trying to start from the drawing board, and perhaps their preferences leaned towards the Democrats; while the those moving into these states from the Union base were likely to be Republican supporters.

In fact, the proportion of mixing of the Northerners and Southerners is reflected in the pattern of the Republican and Democratic representation in these states as well. For example, the Republican strength seems to be highest in the Canadian

borders (where Northerners presumably settled in greater proportions than the Southerners) and falls gradually downwards while that of the Democrats seems to increase towards the South (see Keyssar [49], for a lively description of these ideas).

There have been various attempts at measuring the degree of inter-party competition at the state levels (see Golembiewski [39], Schlesinger [77], Ranney and Kendall [70], and Jewell [48], for example). Some authors (Ranney and Kendall) look at competition by looking at percentages of votes won by the two parties in elections for three offices –President of the United States, United States senator, and governor from 1914 to 1952. Some authors (like Schlesinger) have added another dimension to counting votes - that of how long an incumbent keeps office, or alternatively how often does power change and have then classified inter-party competition in various states taking both dimensions into account. He considers election results from 1872 to 1950 for both gubernatorial and presidential elections. Another approach (suggested by Malcolm Jewell, see Munger [66]) focuses on interaction of the executive and legislature in policy making. No matter what the measures have been, the conclusions have been very similar -

> The Western Mountain States thus appear to have a higher degree of political volatility than is to be found elsewhere (Schlesinger).

And a picture (refer to figure 2.3 in chapter 2) marking the competitive states (according to Schlesinger) and the suffrage states makes the story even more convincing.

Moreover the progressive movement at that time also seem to have gathered substantial momentum in the western states. Keyssar [49] in fact talks about the strength of the People's party being crucial to the 1893 success of women's suffrage in Colorado. He says the following-

> What seems to have tipped the balance in a handful of western states... was a combination of several additional ingredients. One was a more fluid pattern of party competition, due in part to the strength of the insurgent Farmers' Alliance and shortly later, the People's Party.

The Populist Party gradually called the Peoples' Party was born as the Farmers' Alliance out of agrarian revolt that rose out of falling agricultural prices after the panic of 1873. The Alliance also developed political agenda besides economic goals, primarily the demand for silver coinage to counter the gold standard, believed to be causing deflation in agricultural prices. The Populists' were the first political party to actively include women in their affairs.

Hence given this political background in the western states, we now turn to the partisan competition hypothesis capturing the incentives for extension of franchise in the face of close partisan competition and part of the population (young males, progressives) wanting such extension.

3.2.1 Partisan Competition Hypothesis

In the related theoretical model, enfranchisement is modeled as a response to political incentives within the economy. It is a two-period model of elections where in the first period the issue to be decided is whether or not to extend franchise, whereas in the second period some other policy is under consideration. In period 1, only part of the population is enfranchised. However, they may decide to extend the franchise before policy decision is made in the next period. There are two parties, A and B. Each proposes an enfranchisement platform; then the voters vote on the one prefer. After that, there is a new election in the next period, period 2, where policy is decided.

Consider extension by the parties under close competition (which means both the parties have lots of partisans) in equilibrium. Here, to begin with, policies that

parties propose are expected to be close to their favorites but not exactly their favorites, because not being strongholds, they also have to pay attention to what the other party is proposing and what the electorate is wanting. The electorate understands this, being forward looking rational voters, and part of the electorate (progressives, for example) wants to influence the future policy proposals of the parties to get the possible equilibrium outcome as close to their own favorites as possible.

Extension of the electorate helps them do this. For example, extension of suffrage to women would have helped (male) progressives get closer to their favorite (future) outcomes like prohibition and anti-child labor laws. And to please this part of the electorate and increase their current probabilities of winning (which is also important to the parties given high rents from office), parties propose extension, even if that means moving away from their favorite platforms in future. In our case, this means that even the Democrats (who were less pro-suffrage than the Republicans) would propose suffrage extension. Republicans also propose extension, but they're not hurt since they move towards their favorite policy outcome by extension (they were pro-suffrage).

Now consider when each of the parties has a stronghold (like the Democrats in the South or the Republicans in the North). In this case, irrespective of the affiliation of the stronghold party, all the incentives described above break down, and no extension arises in equilibrium. The stronger party announces its favorite policy (and wins for sure almost), in which case, the best response of the smaller party is also to announce its favorite (even though it will most probably never be implemented). Hence the 'progressive-minded' electorate knows that they have no hope of influencing any future policy outcomes. And the parties have no incentives to please the current electorate either to win, since winner is determined by overwhelming partisan support alone, irrespective of policy proposals.

Hence the empirical predictions of this theoretical model would be very high chances of suffrage when there is close competition and very small chances of suffrage when either party is a stronghold. See chapter 2 of this dissertation for further elaboration of this model.

3.2.2 Testing Empirical Predictions of the Hypothesis

Of course, it is difficult to determine the distribution of partisans in the population so we take the representation of the parties in state legislatures to reflect the preferential divide of the electorate. However, notice that this confounds the two different kinds of preferences that an individual might have for a party - partisan preference and the preference for the policy being decided. (See the section on checking robustness to find how 'stability' of the legislature can be used to infer about the kind of underlying preferences the electorate might have.)

So drawing from both literature and theory we come up with the following political hypothesis[3]:

---

[3]One of the discernible political features of the Western states is the dominance of the federal authority – firstly large proportion of land was owned by the government which could therefore be used for extracting minerals, using timber, grasses etc. under federal permit, and secondly, most of the Rocky Mountain states were poor and heavy federal investment was required for any major public works undertaking, like irrigation for agriculture, generation of hydroelectric power etc. Thus federal action constitutes an important part of political life. This probably contributes to large 'rents from office' in the western states. Shefter [79] , writes regarding characteristics of politics in Western states,

> ...in efforts to gain office and remain in power, politicians could draw on the resources of 'the company' - the railroad or mining corporation that overshadowed all other local institutions.

High rents from office also play an important role in the theoretical model. So another hypothesis to test would be

1. *The higher the rents from office, the higher the probability.*

But we leave this for future work.

Also, the progressives were more prominent in the west that in the rest of the country so we could test a hypothesis like

2. *The higher the vote share of the 'third' party, the higher the probability.*

1. *The probability of suffrage increases when the seat shares of the Republicans and the Democrats are close (in both the houses).*

Since the legislature consists of two houses, 'closeness' of seat shares in both the houses perhaps would imply overall competition. Hence, though the theoretical model does not distinguish between upper and lower houses, 'an even distribution of partisans in the population' would perhaps translate into 'close seat shares of the Republicans and the Democrats in both houses of the legislature' (this is more elaborately justified in the 'Data' section of the chapter where the exact specification of the 'closeness' variable is proposed).

3.3 Other Hypotheses

Now let me revisit some previous works and ideas in order to formulate some testable hypotheses.

3.3.1 Demographic Factors

> I think there's just one kind of folks. Folks.
> -To Kill a Mockingbird by Harper Lee

Even though the west was exceptionally receptive of women's suffrage, the rest of the country seemed sluggish to come to grips with it. In the South, the demographics seemed to have aided such aversion. Firstly, the predominantly agricultural rural social structure delayed the expansion of the urban, educated, professional middle-class - the class most supportive of women's suffrage.

Secondly, the southerners feared that allowing franchise to women would inevitably lead to demands of franchise for the blacks, something imperceptible and to be resisted with all might. (Moreover the Democrats in the South were

---

In fact, we did include the share of the 'third' party in empirical tests, while checking for robustness later on in this chapter but never did it turn out to be significant.

repellent to the idea of a national amendment, and perceived it to be yet another federal infringement on states' rights and hence their opposition to women's suffrage.)

Since the western states were among the first to grant suffrage and many of them were mining states with an unfavorable female-male ratio, a reason, often cited is that suffrage was granted to attract women to those states (see Keyssar [49] for a description of all these ideas). Keyssar [49] notes (page 196):

> ... Western states tended to be dominated by land-owning farm families yet included a highly visible number of working-class transients who labored in mining, railroading, and agriculture. Since the latter group consisted overwhelmingly of single males, the enfranchisement of women offered discernible political benefits to the settler population...

Thus we have the following hypotheses:

1. *The larger the proportion of blacks, the smaller the probability (of women's suffrage).*

2. *The larger the urban population, the higher the probability. (This also appears as one of the social hypotheses in the next subsection.)*

3. *The lower the female-male sex ratio, the higher the probability.*

3.3.2 Social and Economic Factors

> I hate to think I've got to grow up, and be Miss March, and wear long gowns, and look as prim as a China-aster! It's bad enough to be a girl, anyway, when I like boys' games and work and manners! I can't get over my disappointment in not being a boy; and it's worse than ever now, for I'm dying to go and fight with papa, and I can only stay at home and knit, like a poky old woman!
>
> -'Little Women' (1868) by Louisa May Alcott

As Jo rues in 'Little Women', we all become conscious of times when the geography of life was strictly gendered and trespassing into the opposite territory

was neither welcome nor dared. Among other things in life, participation in politics in general, and voting, in particular, were strictly considered belonging to the male domain. We will not delve deep here into how the seeds of change bore fruit in the minds of people so that the boundary got blurred and women came to be on an equal footing with men, but that history is also worth exploring (see Scott and Scott [78] for example).

In the decades following the civil war (1865-1869), the mental and psychological matrix of the society was undergoing a transformation. The antislavery movement, no doubt, laid the groundwork for such shifting paradigm, to some extent. Demand for gender equality followed demands for racial equality closely. These decades also witnessed a rise in the urban and quasi-urban middle-class, more receptive towards expansion of civil and economic rights.

It is possible that these social changes were aided by certain technological innovations, furthering the image of women away from just a family member to a more active participant in society. Development of gas lighting, municipal water systems, domestic plumbing, canning, commercial production of ice, improvement of furnaces, stoves etc., helped many women to escape from everyday domestic chores to some extent. College and professional women were increasing on one hand, while unskilled immigrant women joined as cooks and nursemaids in large numbers, on the other. The number of women in paid labor force, increased considerably.

Some of the earlier empirical works like McCammon and Campbell [59] also find change in 'gendered opportunities' playing a significant role in bringing about suffrage. They measure 'gendered opportunities' with three variables that were supposed to capture women's inroads to traditionally male arenas of activity (McCammon and Campbell [59]). These are the number of women who were physicians and lawyers, percentage of all college and university students who

were female and the number of prominent women's organization in states. In fact King, Cornwall and Dahlin [51], find percentage of women working in non-agricultural activities as the only robust variable correlated with suffrage. Hence it seems reasonable to include some measure of women working in my empirical work.

So the possible hypotheses emanating from the social developments at that time are as follows[4]:

1. *The larger the proportion of urban population (which is likely to capture the educated, professional, progressive-minded middle class people), the higher the probability.*

[4]The issues of temperance and prohibition were also significant during this time. Prohibition movement had its origins in the 1840's, mainly spearheaded by pietistic religious denominations like (Lutheranians and Methodists) and it gathered momentum around the late 1880's with the rise of Woman's Christian Temperance Union. Now, though the issues of temperance and prohibition, and suffrage are quite distinct (the former pertaining to personal abstinence, espousing traditional family values, while the latter was part of a more radical feminist movement), women suffrage was invariably perceived as adding 'dry votes' to the electorate. One of the reasons for this may have been support of the temperance leaders to the suffrage issue (suffrage as a means of restricting the social wrongs of men by imposing governmental restrictions etc.).

The effect of this identification of suffrage with temperance was ambiguous as far as suffrage was concerned. On one hand, it could be that the population who espoused temperance might espouse suffrage (the Methodists, Lutheranians etc.), and on the other, they might not (given the temperance was based more on traditional family kind of appeal). In fact, McDonagh and Price [60] find that powerful German opposition to prohibition (both German Catholics and Protestants) spilled over as strong opposition to suffrage as well while Methodist and Presbyterian support for prohibition does not convert into support for suffrage. For sure however, the suffrage movement incurred the opposition of the liquor and business interests of the states (the former fearing prohibition, the latter apprehending unfavorable child labor laws etc.).

On the whole however, it seems this identification with prohibition hurt suffrage more than it helped - it has been found empirically that prohibition is negatively related to suffrage turnout (see McDonagh and Price [60]). However, ideally we would like to test what impact temperance movement would have on the probability of suffrage i.e. whether greater proportions of such religious people (that would have helped temperance movement) aided or abated suffrage. So possible hypotheses might be:

The greater the strengths of prohibition and temperance movements, the smaller(?) the probability.

The greater the proportion of foreign-born immigrants (especially German and Irish), the smaller the probability.

However, in this study we have not incorporated these hypotheses and leave it for future work.

2. *The larger the proportion of working women (mainly in the non-agricultural occupations), the higher the probability.*

3.3.3 Diffusion

> No longer do we speak about "here *or* there"; in the quantum world we speak about "here *and* there"... The most perplexing phenomenon in the bizarre world of the quantum is the effect of *entanglement*. Two particles... are mysteriously linked together. Whatever happens to one of them... causes a change in the other one.
>
> -'Entanglement' by Amir D. Aczel

What is true of particles in the world of quantum physics (world of small particles) seems to be equally true for something as different as behavior of states with respect to adopting decisions. Only instead of 'entanglement' it is called 'diffusion'. The idea simply is that one state tends to behave like its neighbors when it comes to adopting policies, ceteris paribus.

With clustering of states in the west, one would naturally be interested in suffrage diffusion. The idea of diffusion is that states adopt a particular policy not only as a response to its internal determinants (demographics, politics etc.), but also by the adoption decisions of neighboring states. In the suffrage context, this happens either because, policy-makers learn by seeing the experiences of its neighbors, or because they fear a 'setback' for their states (lose women to other states) if they don't catch-up, or both.

Moreover there are many empirical papers that study the issue of policy adoption and diffusion. For example, Rincke [71] finds evidence of significant policy diffusion and emulation in the case of establishment of charter schools among California school districts. Fishback and Kantor [35] also find positive 'contagion' effect of adoption of workers' compensation laws in states of the USA. On the other hand, Doyle [33] finds no evidence of any significant diffusion of merit-based student grant programs in states using event history analysis.

Given the concentration of states in the west, there is possibility of diffusion in suffrage enactments and hence I hypothesize the following:

1. *States with neighbors having women's suffrage had higher probability of granting women's suffrage.*

3.4 Empirical Methods

To test the above hypotheses, I do an event history analysis of states granting full suffrage to women[5]. The entry time for all states is taken to be the year 1880 in which all the states were in the risk-set (none of them had any women's suffrage)[6] and we follow each state till 1919 when the data get censored because women's suffrage gets federally mandated in 1920. Obviously, states that did not enact full suffrage by 1919, never left the risk set and therefore have forty years of observations (1880-1919) while those that enacted suffrage within this period left the risk set at the year of enactment and appear in the data until that point. For example, a state like Alabama will have observations for all the forty years while California will have them for thirty one years (1880-1911) having enacted full suffrage for women in 1911.

[5]Notice that in case a state had rescinded suffrage and then granted at a later date (but before the nineteenth amendment), we include the latest date for that state. In duration analysis terminology, we include the date at which the subject (state) leaves the risk set (by adopting suffrage) and does not enter the risk set again. For example, in Utah, suffrage, that existed from 1870 was annulled in 1887 which was again reinstated in 1896. So for this model, we include the date 1896 at which Utah leaves the risk set and doesn't enter again. (I hope to include spells of risk later.)

[6]Some of the territories had granted women's suffrage prior to 1880 and these were included in the data set only after attainment of statehood. For example, Wyoming had enacted full suffrage for women in 1869 as a territory. It became a state in 1890 and had retained full women's suffrage. Hence Wyoming, as a state, enters the risk set in 1890 but exits immediately having women's suffrage already, so that Wyoming appears for only one year 1890. Other jurisdictions in this category are Utah (which became a state in 1896 but had women's suffrage as a territory from 1870). Other places like Idaho, became a state in 1890 and granted women's suffrage as a state in 1896. Hence for Idaho, the window starts in 1890 and it has 7 observations (1890-1896). Likewise the other jurisdictions whose windows start later than 1880 (because they become states later) are Montana, North Dakota, South Dakota, Washington and Oklahoma.

Here we look at discrete-time methods of event history analysis. First of all, only sixteen states (out of forty eight) had granted full suffrage over a span of forty years, with ties (more than one state adopting women's suffrage in a given year) occurring on five occasions, so that the nature of the data has a discrete flavor. Secondly, though a legislature can presumably adopt a policy anytime within a legislative session, the issue of interest to us is not knowing exactly when adoption occurred within a legislative session, but rather when adoption occurred relative to other states. In such analysis, the year in which a policy was adopted is sufficient to mark the occurrence of an event. Hence while policy-adoption may be a continuous process in principle (and we could easily discern exactly when change occurred from vote recordings in state legislatures), a discrete-time model is possibly better suited for our topic.

Given this, we use the complementary logarithmic function to estimate the coefficients. The reason for this is that firstly, it is the discrete-time analog of the continuous-time Cox proportional hazards model, for which no parameterization of the baseline hazard is required (i.e. no duration dependency needs to be specified). And secondly, this function is asymmetric (unlike logit and probit) with fatter tail towards 0 and hence is more suitable for rare event discrete-time EHA, like in our case. This function looks like follows:

$$\lambda_{i,t} = 1 - e^{-e^{x'_{i,t}\beta}}$$

where $\lambda_{i,t}$ is the probability that state i grants full women's suffrage in time t (and doesn't annul it before 1920) and $x_{i,t}$ are the covariates.

3.4.1 Hazard and Survival:

Before including covariates in our analysis, it is helpful to look at the probability of a state granting full suffrage merely as a function of time.

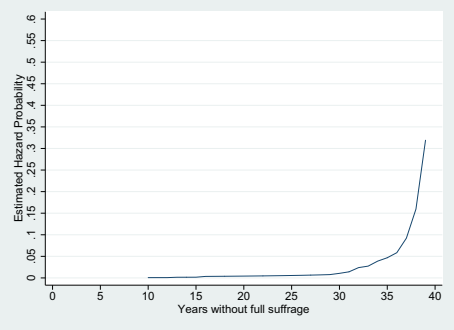


Figure 3.1: Hazard probabilities

As the graph of the hazard probabilities shows, during the forty years under consideration (1880-1919), the probability of full suffrage never reached even 0.5. There was no hazard for the first ten years (1880-1890) with no state having any suffrage during this time. And was very low for the next twenty years (till 1910) with only four states granting suffrage in first thirty years. From 1910 onwards however there is a rise in the probability that increases sharply towards 1917 (with five states granting suffrage in 1917 and 1918 alone).

So, as is evident from the graph, duration dependence seems to be highly non-linear, nearly exponential in this case. A similar picture arises with survival probabilities (probabilities of not enacting full suffrage) as well with survival being very strong till about 1910, then gradually falling for sometime and then rapidly falling.

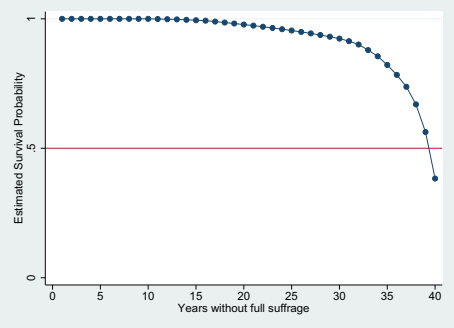


Figure 3.2: Survival probabilities

So it seems, that even without considering any other factor, the probability of granting suffrage began increasing with the passage of time. Hence it might be

insightful to include duration dependence in our analysis.

Correcting the standard errors

Models like logit, probit,or complementary logarithmic functions like the one I use here, require the assumption that observations are independent across time and space which is clearly not the case for policy adoption in states. Several internal determinants of suffrage like political competition, composition and nature of the population etc., are likely to be highly correlated among neighboring states, having similar geographic, climatic, historical and cultural environments. In fact 'similarity' is also cited as a reason for 'diffusion' of policies among neighboring states.

The most familiar remedy to this problem is the use of clustered standard errors, an extension of robust variance estimation. Robust variance estimation allows for the relaxation of the assumption that the error terms are identically distributed, and clustering allows the further relaxation of the assumption of independence between observations in the data. To use clustering, we need to partition the data into groups such that observations are correlated within groups but not across them.

In each of the empirical models therefore, I use the most natural clustering that is clustering by each state (which assumes that the observations of each state is correlated across the years but that this correlation does not persist across states).

3.5 Data

There are two sources from which I have collected the data. The demographic variables have been constructed using data available at Inter-University Consortium for Political and Social Research at the website http://www.icpsr.umich.edu/.

Specifically I have used "Historical, demographic, economic, and social data: the United States, 1790-2000" (study # 2896), especially parts 15 - 24 (containing censuses from 1880-1920). The sources for the ICPSR data are the published volumes of the decennial censuses of the U.S. Bureau of the Census, beginning with the Tenth Decennial Census of the United States (for the year 1880) up to the Fourteenth Decennial Census (for the year 1920).

The composition of the state legislatures are also from ICPSR, available under "Partisan Division of American State Governments, 1834-1985" (study # 16). This data provides information on the number of seats held by the major and minor parties in both houses of the state legislatures (as well as the party identification of the state's governor during the term of each legislature in the United States in the period 1834-1985, which is not used for my present purpose). Data are presented annually and biennially for every legislature. The data from 1834 to 1868 were collected by W. Dean Burnham, Massachusetts Institute of Technology. Data for subsequent years were added by the ICPSR staff.

To check test the hypotheses, I include the following variables in my analysis. The dependent variable is a binary variable that equals 1 for state i in year t if state i has enacted full suffrage for women in year t. It is 0 otherwise. Now consider the independent variables. Let us first consider capturing the political factors. We have the following time-varying political covariates aimed at capturing partisan distributions and competition: percentage of Republican seats in the upper house; percentage of Democratic seats in the upper house; percentage of Republican seats in the lower house; and percentage of Democratic seats in the lower house. Now let closeness of seat shares in one of the houses, say upper house, be defined as follows:

$$\text{closeness in upper house} = 1 - \frac{\text{absolute(\% of Rep seat - \% of Dem seat)}}{\text{\% of Rep seat +\% of Dem seat}}$$

Hence, closeness is a fraction that becomes 0 (there is no closeness) when the sum

and the difference of the seat shares of the two parties are the same (so that the ratio is 1) which happens when one of the parties has no seat share at all (and the other has a stronghold). As the difference in seat shares falls, the ratio falls and closeness rises, with the maximum being reached when the seat shares are exactly equal so that the ratio vanishes and closeness equals 1 (seat shares are very close). Similarly, we define closeness for the lower house. Now competition would be really close in a legislature when the seat shares in both the houses would be close. Hence to capture this overall competition in the simplest possible way, we define 'competition' as an indicator variable as follows:

$$\text{competition} = \begin{cases} 1 & \text{if closeness in both upper and lower houses} > .89 \\ 0 & \text{otherwise} \end{cases}$$

Justification of the competition measure: One can presumably come up with various measures to capture 'competition'. As suggested by the political science literature, these could be 'static' in the sense of capturing competition at every point in time, or 'dynamic' in the sense of capturing competitiveness over time, or a combination of the two. For example, static measures might look at distribution of seat shares (among the major and third parties) in one or both the houses or the legislature as a whole while dynamic measures may try to capture the stability or otherwise of one or both houses in the legislature to infer about competition. I study some of these alternative formulations in 'Checking the Robustness' section. For the main analysis however, I have used the static measure of competition as formulated above and hence it requires some justification.

At the outset, I need to justify such a measure on the basis of its correspondence with my theoretical notion of partisan competition. On a more technical level, there are three things pertaining to the construction of my measure of competition that need to be justified - first, why not use the underlying continuous index

(called closeness) as a measure of competition instead of constructing a dummy based on it; second, why the threshold .89 for the index instead of any other; and third, why consider both the houses (and not any one or the legislature as a whole). Below I provide reasons for my choices:

Correspondence with the theoretical measure of competition: The notion of the distribution of partisanships in the electorate becoming more and more even, is theoretically captured by the height of the distribution becoming smaller and smaller. This entails more and more masses towards the tails of the distribution and less towards the centre. If the centre represents preferences which are 'neutral' (not biased towards one or the other part) and the tails represent partisan preferences for one or the other party, then the distribution becomes such that people are more and more partisans and less and less neutral.

Assuming the composition of the legislature reflects the preferential divide of the electorate, an electorate that is evenly divided among the two parties according to their partisan preferences, is probably captured by even split in the seats of the legislature (in both the houses). (Please refer to section on checking robustness to see the nuances of such an argument.)

Why an indicator variable: Let us start with an illustration of the problem. Consider any one of the houses and suppose the split of 100 seats is 98-2 (yielding a closeness index of .04). And consider the smaller party gaining two more seats. In that case, the split becomes 96-4 (and closeness becomes .08). However, there has not been any significant threat to the majority party and the policies they espouse, in this case. On the other hand, consider the split 52-48 (closeness is .96). Here, if the smaller party gains two seats the balance of power is enormously affected (and so is the probability of enactment of suffrage, for example). Here closeness becomes 1 (changes by .04, like before). Hence it is likley that for the same change

in the continuous variable, the probability changes very discontinuously and non-linearly. Now, in our chosen model (which is a non-linear model), change in probability is different for different values of the independent variable (and hence changes non-linearly) but we can show that the change is essentially bounded and hence not suitable for capturing the possibly large discontinuous jumps in suffrage probabilities.

More rigorously, for small changes in x (assume just one independent variable for expositional ease), changes in the probability of suffrage, $(\frac{\partial \lambda}{\partial x})$, is bounded (the latter not only contains a constant slope parameter but also some non-linear transformation of x that is bounded). Hence change in probability is smooth and bounded, unsuitable to capture the kind of story that we have in mind. Hence the discrete jumps in probability can better be captured by defining a dummy so that probability can suddenly increase when the dummy turns from 0 to 1 (competition increases beyond a threshold for the parties to feel the pressure).

Why .89**:** To define the indicator variable I use the cut-off .89 (for both the houses), which corresponds to the 90^{th} percentile of the index of competition for the Lower House in the sample. This translates to about 4 to 5 seats (on average) out of 35 (on average) in the upper house (the data show most upper houses having about 35 total seats) and about 12 to 13 seats on average out of 120 in the lower house (again, most lower houses having about 120 total seats). So we are looking at a split of about 15 - 20 in the upper house and a split of about 66 - 54 in the lower house.

The 90^{th} percentile for the competition index of the Upper House is about .85. First of all, a difference of 15 seats out of 100 may not be reflecting a close balance of power. Moreover, though a higher threshold would have corresponded with higher and higher percentile for both the houses, it would leave out more and more data points. So I settled for the higher competition index corresponding to

the lower percentile, in order to capture maximum possible competition for the largest number of states[7]. (It is implicit in the above argument that I wanted to choose the same threshold for both the houses so as to not bias the measure for or against one of the houses.)

Why both the houses: Consider a simple but extreme example for illustration. Suppose each of the houses has 100 seats and out of the two parties, one party holds all the seats in one of the houses, while the other holds all the seats in the other house. Then out of 200 total seats in the legislature, each party has 100, but of course, there will be no competition in the sense of close balance of power between the parties, since in each house one of the parties is a clear majority. However if the parties were split 50-50 in each house, then again, the total would be a split of 100-100 and would also imply close competition since power would be really balanced.

Hence, close competition in both the houses individually is sufficient for overall close competition but overall close competition is not so. Therefore, I look at each house separately to be sure that I capture close balance in each of the houses (implying close balance overall).

Continuing with data description, there were about 30% missing observations for the political variables, namely the seat shares of the two parties in the two houses. This however confounds missing data i.e. no data for years in which the legislature did meet, and years in which the legislature did not meet. Moreover most of the missing data were in the early years of our window (around 1880's). So I tried running the analyses in three ways to better understand how big this problem of missing data was. Either I used the values of the last available year to substitute

[7]Notice that since the Upper House is a much smaller house (about 35 seats on average), changing the threshold from .85 to .89 will mean a difference of about 1 or 2 seats.

a missing value or I linearly interpolated between the two nearest available values or I dropped all the missing observations altogether, while estimating the models.

It turns out that the results of the first two methods (keeping the last available value and linearly interpolating between nearest values), yield very similar results. And dropping all the missing values improves the results in terms of log-lokelihood and p-values of estimates (please refer to the section titled Women's Suffrage-A in the Appendix for the results).

The proper way to do this of course, would be to figure out which are the missing data and which are simply years in which the legislature did not meet. A state is only 'at risk' to extend suffrage in a year in which the legislature met. So one method would be to include only those years in the dataset. But it is likely that the data has observations on the composition of legislatures for years when the legislature did not meet, so that this would mean dropping many other observations.

Again, linear interpolation could influence the composition of the legislature by either making it reach the cut-off of .89 (the value that is relevant for our measure) too soon or too slowly. In any case, therefore, the results presented here (which are very close to the results with linearly interpolated data and somewhat close to those with the data where missing values are dropped) are those where missing observations have been substituted with the last available ones.

Next I have a set of demographic time-varying covariates like percentage of urban population, sex-ratio (expressed in percentage to make it compatible with the unit of measurement of the other variables) and percentage of blacks, corresponding to our hypotheses. (Notice that some variables are highly correlated with one another like percentage of blacks highly negatively correlated with the percentage of illiterate and the percentage of urban people highly positively correlated with

percentage in non-agricultural population.) In the results presented, I just include the only variable that has been found robust in women's suffrage studies so far, namely, percentage of women working in non-agricultural occupations.

Since the data for these variables come from decennial censuses and hence are available every ten years, in this case I linearly interpolate the values of the variables for all the intermediate years.

I capture 'diffusion', in the simplest way, by defining a dummy which takes the value 1 for a state in a given year, if at least one of its neighbors (the states that share a geographic boundary with this state) had adopted suffrage prior to that year. It is 0 otherwise.

I also include duration dependence linearly by including the number of years from the beginning of the statehood or the beginning of the window (1880), whichever is later, that a state has not enacted suffrage.

3.6 Analysis and Results

Before presenting the regression results, it will be helpful to have a look at the descriptive statistics of the chosen variables.

Table 3.2: Descriptive statistics

Variable	Mean	Standard Deviation
Closeness of seat shares (Upper House)	.4162	.3000
Closeness of seat shares (Lower House)	.4636	.3066
Competition dummy	.0234	.1513
Sex-ratio (female/male) in %	92.5668	10.0997
% of urban population	31.8544	21.4824
% of black population	10.7270	16.1619
% of female working	14.6087	6.0971
Diffusion dummy	.1471	.3544
Duration dependence	19.5946	11.3462
Number of observations	1665	

Coming to the regression analysis, I estimate three different models. First of all, standard errors in all the models are 'clustered' by states, since errors are very likely to be correlated within a state and over time. Given clustered errors in all models, model (1) is the basic model which does not include the diffusion dummy, nor considers duration dependence. Model (2) adds the diffusion dummy to model (1). Model (3) adds possible duration dependence to model (2). The main regression results are summarized in the following three tables as follows:

Table 3.3: Basic model

Model 1[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Competition dummy	1.2151	1.1561	.293	.0069
Sex-ratio (female/male) ***	−.0624	.0154	.000	−.0002
% of urban population*	.0466	.0244	.056	.0001
% of black population	−.0906	.0577	.116	−.0003
% of female working	−.1040	.0836	.214	−.0003
Constant	.9519	1.2383	.442	
Log likelihood	−78.8032			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. The marginal effect of a continuous variable, x, on the dependent variable, y, is computed as the slope $\frac{\partial y}{\partial x}$, at the sample mean of x and holding all the other variables constant at either their sample means (or other specific values). The marginal effects of the dummy variables are based on switches from zero to one, holding all else constant at sample means (or other specific values). The probability of suffrage for model (1) was computed at the sample means of all the variables and it is about .003. Marginal effects are also calculated at this probability.

Table 3.4: Model with 'diffusion'

Model 2[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Competition dummy	1.8526	1.1488	.107	.0171
Sex-ratio (female/male)***	−.0559	.0204	.006	−.0002
% of urban population***	.0847	.0231	.000	.0003
% of black population	−.0526	.0557	.345	−.0002
% of female working**	−.1998	.0982	.042	−.0007
'Diffusion' dummy***	2.8210	.5460	.000	.0142
Constant	−1.0820	2.0234	.593	
Log likelihood	−66.8806			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for model (2) by letting the 'diffusion dummy' equal .463 (not equal to its sample mean of .147) and holding others at their sample means.

Table 3.5: Model with 'diffusion' and 'duration dependence'

Model 3[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Competition dummy*	2.2270	1.2755	.081	.0228
Sex-ratio (female/male) ***	−.1059	.0339	.002	−.0003
% of urban population***	.1277	.0295	.000	.0004
% of black population*	−.1704	.1036	.100	−.0005
% of female working***	−.4792	.1559	.002	−.0014
'Diffusion' dummy***	2.0900	.6606	.002	.0151
Duration dependence**	.1215	.0524	.020	.0004
Constant	3.2771	2.9755	.271	
Log likelihood	−61.5141			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability, like in model (1) of about .003 (at which the marginal effects would be calculated) was computed for model (3) by letting 'duration dependence' equal 42 (not equal to its sample mean of 19.5946) and holding others at their sample means.

The marginal effect is measured by $\frac{dy}{dx}$ where y is the dependent variable and x is the continuous independent variable, at the sample mean of x, and all the other variables are held constant at their sample means. The marginal effects of the dummy variables are based on switches from zero to one, holding all else constant at sample means (or other specific values).

Before analyzing the results, notice that for continuous variables, the marginal effects seem to be very small relative to those of the dummies. It is likely that the marginal effects of the continuous variables are downward biased compared to

those of the dummies since, for the former type of variable, the marginal effects measure change in probability for infinitesimal change in the variable, while for the latter, the marginal effects capture the change in probability for (relatively) large discrete jumps in the underlying variable. Hence to create a comparable ground for the effects of the two types of variables, we present the elasticities for the continuous variables. Recall elasticity in this case will measure the percentage change in probability (y) for a 1% change in the value of the continuous variable (x). Mathematically, it measures the percentage change in y relative to that in x, i.e. $\frac{\frac{dy}{y}}{\frac{dx}{x}}$ i.e. $\frac{dy}{dx} * \frac{x}{y}$ or $\frac{d(lny)}{d(lnx)}$. In table 3.6, the elasticities of the continuous variables are presented.

Table 3.6: Comparing elasticities of continuous variables

Elasticities[a]	Model 1	Model 2	Model 3
Sex-ratio (female/male)	−5.7720	−5.1671	−9.7911
% of urban population	1.4826	2.6939	4.0612
% of black population	−.9703	−.5628	−1.8257
% of female working	−1.5166	−2.9144	−6.9897

[a]To facilitate comparison across models, the elasticities were computed for the same probability of about .003 in all the models by choosing the 'diffusion' and 'duration dependence' variables in models (2) and (3) suitably, as described in the footnotes of tables 3.4 and 3.5.

Political Findings:

All the models suggest significant roles of the political players. Even in the basic model, competition, which means close seat shares between the Republicans and the Democrats in both the houses, seems to be affecting suffrage positively. And given the legislature is assumed to be representative of the electoral divide, it means that close partisan distributions in the population would have proven conducive to suffrage enactments.

Looking at the coefficient of competition in the three models, it increases as diffusion is added (from model (1) to (2)), which suggests that not only is competition robust to inclusion of diffusion, but also its influence becomes more prominent when other factors affecting suffrage are factored out. And in fact, in model (3), when possible duration dependence is added, the coefficient of competition not only increases in magnitude (from that in model (2)), but also becomes significant at 90%.

Consider the column of 'Marginal Effect' in all the three tables. Since competition is a dummy variable, the marginal effect (or the economic impact) measures the change in probability of full suffrage when competition changes from 0 to 1, i.e., when closeness of seats shares in both the houses crosses the threshold .89, and other variables are held constant at their means. Consider model (1) first. Here we find that probability is likely to increase by .69 percentage points when competition changes from 0 to 1.

Comparing the marginal effects across models, notice that for a given probability of suffrage (like that in model (1)), the marginal effect of competition increases (as suggested by the regression result) as diffusion and duration dependence are added. More specifically, the marginal effect of competition is likely to be about 1.7 percentage points when only diffusion is added (model (2)) and about 2.3 percentage points when both diffusion and duration dependence are added.

An intuitive explanation for this might be as follows: It is possible that some states were enacting women's suffrage just because some of their neighbors were enacting it (diffusion), even when competition was low. In the absence, of a diffusion variable, this could weaken the strength of the competition variable (since it would seem that women's suffrage can happen even without strong competition). However, once the effect of diffusion is taken care of, the effect of competition only stands out more clearly.

For example, once diffusion is included (model (2)), one finds that the economic impact of competition is 1.7 percentage points, compared to .02 percentage points of sex-ratio, 1.4 percentage points of 'diffusion', .03 percentage points of percentage of urban population, .02 percentage points of percentage of black population and .07 percentage points of the percentage of women working in non-agricultural occupations.

The result is similar for model (3) where competition affects probability of suffrage by 2.3 percentage points (which is higher than the impacts of the other included variables).

Demographic and Social Findings:

Among our other findings, sex-ratio and percentage of urbanization turn out to be consistently significant and in the popularly believed way. That is, the coefficient of sex-ratio (female/male) is negative implying the worse the ratio (like in the western states) the higher the probability of suffrage while that of percentage of urban population is positive implying the higher urban population conducive to suffrage. Of course, percentage of urban population is representative of other social and demographic phenomena like larger middle-class, larger professionals who have typically been found to espouse suffrage.

Percentage of blacks doesn't turn out to be significant in the first two models but once both diffusion and duration dependence are allowed (as in model (3)), the coefficient becomes significant and in the expected way by negatively affecting full suffrage enactments.

What is a little surprising is that percentage of women working in non-agricultural occupations turn out to be significant from model (2) onwards but in a way opposite to that expected. The coefficient is negative. However the sign does

seem plausible given the low percentage of working women in the western states, the part where suffrage was granted earliest.

To see the effect of these variables on the probability of suffrage, we look at their elasticities. Recall elasticity in this case will measure the percentage change in probability (y) for a 1% change in the value of the continuous variable (x). Using this we find the following for model (1): for a 1% worsening of sex-ratio, the probability of suffrage is likely to increase by about 5.8%; for a 1% increase in the percentage of urban population, the probability of suffrage goes up by about 1.5%; for a 1% worsening of percentage in black population, the probability of suffrage increases by about .97%; and for a 1% decrease in percentage of female working in non-agricultural population, the probability of suffrage is likely to go up by 1.5%.

Hence it seems that the probability of suffrage was very responsive to worsening of the sex-ratio (very elastic), somewhat responsive (elasticity is slightly greater than 1) to increases in urbanization and worsening of female participation in non-agricultural occupations, and not quite responsive to changes in the black population (elasticity less than 1).

The nature of the marginal effects and elasticities of these demographic and social variables are qualitatively similar for models (2) and (3) - sex-ratio has the highest effect on the probability of full suffrage among the others.

Comparing across models, however, one finds that for a given probability of suffrage, the elasticity (and marginal effect) of sex-ratio falls when diffusion is added (thereby comparing between models (1) and (2)). This could mean that the effect that was being captured by sex-ratio was in fact partly being caused by diffusion. Hence inclusion of diffusion as a separate variable takes away that part of sex-ratio that was confounded with diffusion, and the effect of the former falls.

For example, it could be that some states were enacting women's suffrage just because some of their neighbors were enacting it (diffusion), and not because their sex-ratio was not in favor of women. In the absence of a diffusion variable, this could strenghten the sex-ratio variable (since it would seem that adverse sex-ratio was driving women's suffrage enactments, even though it was just emulation of neighbors). However, once the effect of diffusion is taken care of, the effect of sex-ratio only stands out more clearly (falls in this case).

However when both diffusion and duration dependence are added (model (3)), the elasticity and marginal effect of sex-ratio rises again and is larger than in both the other models. It is possible that some states without adverse sex-ratio were enacting women's suffrage just because they thought it was inevitable with the passage of time (duration dependence). In the absence, of a duration dependence variable, this could weaken the strength of the sex-ratio variable (since it would seem that women's suffrage can happen even without adverse sex-ratio, just with time). However, once the effect of duration dependence is taken care of, the effect of sex-ratio stands out more clearly (rises in this case).

The marginal effect of percentage blacks also work in a similar pattern (like sex-ratio) across the models, and presumably can be interpreted in the same way. However, the marginal effects of the percentage of urban population and percentage of female working in non-agricultural occupations, increase as diffusion and duration dependence are added, which can be explained as above.

Time and Diffusion:

From the regression results, both diffusion and duration dependence seem to have played very positive and significant roles in bringing about full suffrage. Hence states were more likely to enact suffrage if neighbors had done so and if it had been a long while without suffrage. The marginal effect of diffusion is 1.4 percentage

points (model (2)) while that of the number of years without suffrage is about .04 (model (3)).

Neighbors' decisions could have affected policy-makers because states might have learned by seeing the experiences of its neighbors, or because they feared a 'set-back' for their states if they don't catch-up, or both. In the women's suffrage context, especially in the west which was sparsely-female-populated, such laws could be the result of 'catching-up' with the neighbors that already had such suffrage laws. The idea would be to make their own states equally attractive for women to come and settle in and not fall behind in this respect to their neighbors.

Also the probability of suffrage went up (although slightly), the longer a state remained without such laws. This could be because of a sense of ultimate inevitability of passage of such laws in the face of widespread gender equality movements, leading to enactment. This sense most likely went up with the passage of time, during which possibly more and more countries and other American states would have enacted women's suffrage. And hence positive duration dependence could very well have foreshadowed the nineteenth amendment that federally mandated women's suffrage.

Comparing the marginal effect of diffusion across models (2) and (3), one finds that it increases (from 1.4 to about 1.5 percentage points) when duration dependence is added. This could be because states could be enacting women's suffrage even without having neighboring states that had already enacted it, just out of the sense that it was going to happen sooner or later. This could lead to weakening the of the impact of 'diffusion' (since it would seem that states without 'diffusion' still had women's suffrage). However, once 'duration dependence' is taken care of, the effect of 'diffusion' only stands out more clearly (rises in this case).

3.7 Checking for Robustness

In this section I consider other plausible measures of competition (always in model (3) that includes diffusion and duration dependence), some drawn from related literature, and some proposed independently. Let us first consider those that are based on composition of the legislature at every point in time (and hence are called the 'static measures') and then we will consider measures that consider the change in the composition of the legislature over time (and hence are called the 'dynamic measures').

3.7.1 Static Measures

Majority Surplus: Smith and Fridkin [81] have used the notion of 'majority surplus' to capture interparty competition in the legislature. The measure basically pools the seat shares of the parties in both the houses of the legislature and for the party who has the majority (share $> 50\%$), it measures the 'surplus' or the number of seats exceeding 50%. And then lower is the 'surplus', the greater is the 'competition'. I included such a variable in my regressions but it was never significant. Moreover the marginal effect is very small.

Table 3.7: Model with 'majority surplus'

Model 3 with Majority Surplus[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Majority Surplus	.0055	.0148	.711	1.87×10^{-5}
Sex-ratio (female/male)***	−.0982	.0346	.004	−.0003
% of urban population***	.1238	.0294	.000	.0004
% of black population*	−.1678	.0991	.090	−.0006
% of female working***	−.4626	.1450	.001	−.0016
'Diffusion' dummy***	1.9290	.6261	.002	.0149
Duration dependence**	.1144	.0506	.024	.0004
Constant	2.8016	2.6030	.282	
Log likelihood	−62.7033			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for this model by letting 'duration dependence' equal 42 (not equal to its sample mean of 19.5946) and holding others at their sample means.

Notice that this measure suffers from the usual drawback of a continuous measure trying to capture the pressure of competition. For example, according to this measure, 20% seats above 50% (i.e. 70% seat share) for the majority party would mean more competition than 30% seats above 50% (i.e. 80% seat share) of the majority party, though both would be situations of very less threat from the opposition and hence almost no competition at all.

Moreover, the measure pools the seat share of both the houses and hence confounds close seat share of each of the houses (that would mean competition) with closeness of seats in the overall legislature (which does not necessarily mean

competition).

Third Party: Smith and Fridkin [81] also include (overall) seat share of the third party in the legislature. Like in their case of delegation of direct democracy in the western states, in the case of women's suffrage too, many believed that the presence of third party, which was mostly the Progressive Party, was crucial in passing reforms like voting rights for women. Hence I included the seat share of the third party in the regression (for each house). Though the sign of the coefficients is as expected (positive), they are not significant[8]. Moreover the marginal effects are very small.

[8]I included third party presence along with my original competition dummy but still there has been no change in the significance of the third party.

Table 3.8: Model with 'third party'

Model 3 with Third Party[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Third party in the Upper House	.0222	.0209	.288	.0001
Third party in the Lower House	.0075	.0060	.212	2.1×10^{-5}
Sex-ratio (female/male)***	−.0969	.0367	.008	−.0003
% of urban population***	.1424	.0328	.000	.0004
% of black population*	−.2650	.1398	.058	−.0007
% of female working***	−.5206	.1580	.001	−.0015
'Diffusion' dummy***	1.9415	.5800	.001	.0126
Duration dependence**	.1272	.0492	.010	.0004
Constant	2.0378	3.3545	.544	
Log likelihood	−61.3492			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for this model by letting 'duration dependence' equal 46 (not equal to its sample mean of 19.5946) and holding others at their sample means.

Split Legislature: Often competition is perceived to be strong when the legislature is split between the parties (that is, one of the parties has a majority in one of the houses, and the other party has it in the other). However, a dummy of identifying split legislatures does not turn out to be significant[9]. And the

[9]If, this dummy of the other hand is interacted with the dummy of competition (close seat shares in each of the houses), the interaction term drops out of the regression for predicting failure perfectly. So it does seem to be the case that a split legislature where the parties are close would be very likely to grant women's suffrage. And this is also very plausible since competition would not be perceived to be strong in case a legislature is split, but the party

marginal effect is about .8 percentage points.

Table 3.9: Model with 'split legislature'

Model 3 with a Split Legislature[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Split legislature	1.3230	1.3459	.326	.0087
Sex-ratio (female/male)**	−.0948	.0391	.015	−.0003
% of urban population***	.1277	.0331	.000	.0004
% of black population	−.2068	.1407	.142	−.0007
% of female working***	−.4827	.1567	.002	−.0016
'Diffusion' dummy***	1.9780	.6115	.001	.0151
Duration dependence**	.1226	.0498	.014	.0004
Constant	2.5409	3.0258		
Log likelihood	−62.2331			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for this model by letting 'duration dependence' equal 44 (not equal to its sample mean of 19.5946) and holding others at their sample means.

Notice that among the static measures, the marginal effect of a 'split legislature' is much higher (about .9 percentage points) than those of 'majority surplus' and 'third party' (whose marginal effects are almost 0).

holding the majority in one of the houses is very large. On the other hand, close seat shares alone indicates strong competition, even when the legislature is not split (that is, the same party has a majority in both the houses), as indicated by the regression results earlier where split was not considered.

3.7.2 Dynamic Measures

Dynamic measures are aimed to capture the stability or otherwise of a legislture over time. The following will help to motivate such a measure for the question of women's suffrage.

Empirical motivation for 'stability': The suffragists (or progressives, maybe) tried to woo politicians for their cause. Now instead of trying and persuading all the people of both the parties, the suffragists were likely to try and persuade the people of the party in power. In case of a stronghold party, the party in power is unambiguous. But in case of close seat shares (like in the western states), a party is likely to be considered powerful, if it has held majority for most of the times.

In case, majority has flipped, there is greater political instability and the suffragists must have found it hard to find supporters of their cause or even get the bill introduced. They wouldn't know who to try and convince because they wouldn't know who would be the majority next time. Hence what would be most conducive to suffrage in the long run is a 'stable' house, in the sense that there is not much flipping going on, so that suffragists would exactly know whose support to try and win. Hence I check whether 'stability' mattered for women's suffrage. Here I define 'stability' as a dummy which is 1 if the majority has not flipped even once in the Upper or the Lower House in the last 3 years.

Table 3.10: Model with 'stability'

Model 3 with Stability[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Stability	−.2149	.6120	.725	−.0007
Sex-ratio (female/male)***	−.0972	.0373	.009	−.0003
% of urban population***	.1271	.0313	.000	.0004
% of black population	−.1952	.1267	.123	−.0006
% of female working***	−.4741	.1497	.002	−.0016
'Diffusion' dummy***	1.9729	.5981	.001	.0151
Duration dependence**	.1170	.0524	.025	.0004
Constant	2.9710	2.8137	.291	
Log likelihood	−62.6803			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for this model by letting 'duration dependence' equal 44 (not equal to its sample mean of 19.5946) and holding others at their sample means.

However, results indicate that only 'stability' would not have mattered significantly. So we turn to the following conjecture.

Empirical motivation for 'stability' and 'closeness': A further complication to this idea of 'stability' would be that not only should 'flipping' be low, but the seat shares should be close as well, for the party in turn to pay attention to the suffragists' demands, knowing that their power is closely contested. Otherwise, like in case of a stronghold party (which is likely to entail low 'flipping' also), there will be no incentives to pay any heed to the suffragists' demands and the suffragists would also have had a hard time convincing an overwhelming majority

(and there was no point convincing the rudimentary opposition).

Hence apparently two contradictory forces of competition must have mattered for suffrage. On the one hand, close seat shares would be conducive, while on the rather, certain degree of stability as to who had the majority, would also be helpful. A theoretical motivation of how this idea of both 'stability' and 'closeness' corresponds with my idea of evenly balanced partisan distributions in the electorate, is provided below.

Theoretical motivation for 'stability' and 'closeness': What I assume when I only consider close seat share to capture the partisan preference in the electorate is that an ex-post 50-50 division of seats in the legislature implies that 50% of the voters had been Democratic partisans and the rest 50% Republican partisans. This could be, but is not necessarily true. Ex-post close seat share in the legislature in itself may not reflect an equally balanced (distributed) partisan preference of the electorate.

For example, this could be the result of ex-ante say 50% of the voters being partisans, 25% in favor of each of the parties while the remaining 50% are the non-partisans or neutrals (those who don't have strong party biases). Suppose these neutrals toss coins to decide the party to vote for (assuming similar policy platforms of the parties which is true in our case, so that policy-wise they are equally better-off voting for each of the parties). In that case the resulting vote share might just be split almost 50-50, even when all of the electorate does not have strong partisan preferences.

However what can distinguish the first mechanism (which is what my theory has and the informal suffrage literature stresses - that of many partisans on either side) from the second is 'flipping'. Majorities would be likely to 'flip' often with less partisans (coin-tossing can go either way), while 'flipping' should be rare, given

strong partisanships. Hence a more accurate empirical analog of my theoretical 'evenly balanced partisan preference in the electorate' would be 'close seat shares and less flipping' or 'close seat shares and more stability'.

In the following regression, I test whether, this hypothesis is true. In particular I test whether stability and close seat shares were important. 'Stability' is defined as before and 'closeness' is taken to be the original measure of competition. And the variable 'stability and closeness' is the product of the two which captures stability in both houses of the legislature as well as close seat shares in both of them. The latter turns out to be very significant.

Table 3.11: Model with 'stability and closeness'

Model 3 with Stability and Closeness[a]	Coefficient	Robust Std. Error	p-value	Marginal Effect
Stability and closeness***	3.3801	1.3096	.010	.0764
Sex-ratio (female/male)***	−.1113	.0354	.002	−.0003
% of urban population***	.1277	.0281	.000	.0004
% of black population	−.1592	.1034	.124	−.0005
% of female working***	−.4711	.1467	.001	−.0014
'Diffusion' dummy***	2.0511	.6518	.002	.0142
Duration dependence**	.1262	.0550	.022	.0004
Constant	3.4758	2.9163	.233	
Log likelihood	−60.5183			
Number of observations	1665			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .003 (at which the marginal effects would be calculated) was computed for this model by letting 'duration dependence' equal 41 (not equal to its sample mean of 19.5946) and holding others at their sample means.

It seems that 'stability and closeness' together had a very significant and large effect on women's suffrage enactments. Its marginal effect at about 7.6 percentage points is greater than that of all the other variables. However the marginal effect of only stability is small (about .07). Hence the results seem to support the story of both, close seat shares as well as a certain degree of stability, proving to be conducive to suffrage extensions to women.

Moreover the results of the last regression also lend support to role of some 'lobbying'-type mechanism at work behind women's suffrage. I do not have explicit lobbying in my theoretical model because preferences of the voters are known a priori by the parties from their position in the voter spectrum. And parties extend voting rights today because part of the current electorate (men) wants extension. Hence both the parties increase their chances of winning today by proposing extension and pleasing part of the men who want extension (like the progressive-minded middle-class), even if that means moving away from their favorite policy platform in future. Now in the theoretical model, which of the voters would want extension is determined by their position in the voter spectrum hence the 'want' is trivially conveyed to the parties, without requiring 'lobbying'. However, in real life, the only way for parties to know voter preferences might be to listen to them, requiring some 'lobbying'-type mechanism. In short, parties must know voter preferences before extending. In the theoretical model the mechanism was obvious. In real life, maybe it was 'lobbying'.

To conclude this section we find that none of the static measures to capture competition ('majority surplus', 'third party', and 'split legislature') as well as only 'stability', among the dynamic measures, turn out to be significant and the marginal effects are 0. Notice that none of these capture the notion of pressure of partisan competition that I have in my theory. Moreover, though my theory does not include lobbying, the regression results that support stability and closeness,

make a very strong case for it and hence I would like to include it explicitly in future theoretical extensions of my model.

3.8 Conclusion

In this paper, my mission is two-fold: Firstly, I propose a hypothesis for explaining early suffrage in the western states and secondly, I revisit some of the existing hypotheses (along with the proposed one), within a discrete-time event history framework. I find that partisan competition contributes positively towards suffrage enactment, along with other demographic variables like adverse female-male sex-ratio and higher percentage of urbanization. There is also evidence of significant diffusion effects and duration dependence.

This paper lends support to the predictions of the related theoretical paper, and therefore suggests an important variable to be considered in empirical models of suffrage extension. Since there is nothing specific in the theoretical model for extension of suffrage to imply extension of suffrage to women only, the predictions can be applied to other extensions as well. However, since the theoretical model does not use tools of redistribution (like taxes and transfers) for extension to occur in equilibrium, (which most of the other theoretical papers in the literature use) my model is general enough to incorporate cases of extension where redistributive repercussions may not have been of paramount concern, like the case of women's suffrage.

Chapter 4

On Existence of Equilibrium with Endogenous Income

4.1 Introduction

> The one thing that doesn't abide by majority rule is a person's conscience.
> - Harper Lee in 'To Kill a Mockingbird'

It is interesting to perceive income taxes as equilibrium outcomes of political games played among political parties. The environment is essentially one where political parties or candidates propose income-tax schedules for the whole population with some objective in mind, typically winning the election and collecting some fixed amount of revenue. An individual voter of the electorate then votes for the platform that maximizes his/her own utility. Some rule is employed for determining the winner and his/her announced platform is then implemented (assuming commitment to announcements on the part of the candidates).

It is well-known in the literature that there exists no pure strategy equilibrium tax schedule when incomes are exogenous (that is, when voter-utility does not depend on leisure-labor choices so that labor-supply decisions do not respond to tax announcements). Assuming majority rule for determining winners, it can be shown that for any announced tax policy of one party/candidate, it is always possible to construct an alternative one that wins in majority voting over the first one and collects the desired revenue. The main reason for this is that, hurting

a smaller fraction of the populace (by imposing a little more tax on them) and favoring a larger fraction (by lowering taxes for them) and still collecting the required revenue, always seem to be possible in such games (see Klor [52] for a formal proof of this idea). The problem is analogous in spirit to the problem of cake-sharing among three individuals. It is always possible to take a little bit from one individual and distribute that among the other two and this allocation gets the support of the two beneficiaries who constitute a majority.

In this paper, we study the issue of existence of equilibrium schedule with endogenous income (when people do make labor-leisure choices in response to tax announcements). This could be interesting because perhaps, beyond a point, it will not be possible to collect the same revenue anymore by reducing taxes from somebody and taxing others since the high-taxed section can then change their labor choices, and earn the income of the low-taxed section and thus not pay the taxes that were aimed at them. In other words, when people begin to respond by changing labor-supply to tax announcements, it is perhaps not possible to design alternative incentive-compatible tax schedules and still meet revenue requirements beyond a particular distribution of taxes in the population. In that case, that particular tax schedule is likely to be the equilibrium one, no one would be able to do better by announcing some other platform.

However if we allow for all kinds of tax schedules (both marginally progressive and regresssive ones), our conjecture would be that there would still be no pure strategy equilibrium in the endogenous income case as in the exogenous income one. We show that with exogenous income, there is no pure strategy equilibrium even when tax schedules are restricted to be marginally progressive only, while with endogenous income there can be pure strategy equilibrium. We in fact prove in this paper that if the revenue to be collected is 'high', then there exists an equilibrium and the equilibrium tax schedule is linear (as concave as you can get

in the space of convex schedules); if the revenue belonged to some 'intermediate' range of values then also there exists an equilibrium and the equilibrium schedule is strictly convex. For all lower revenue, there exists no equilibrium.

Let us say a few words to justify our convexity assumption. Marhuenda and Ortuno-Ortin [57] showed that any marginal rate progressive tax will always win in majority voting over any marginal rate regressive tax (in the absence of negative tax) for all positively skewed pre-tax income distributions. This result, however, ignores incentive effects. The voting behavior is therefore to vote for that schedule that taxes them less. T.Mitra et al. [64] introduced "relative standing" concerns of the agents in such a model. Their model incorporates "interdependent preferences" in the following way: "an individual is assumed to vote for a tax function over another whenever the former treats her better than the latter with respect to her absolute and relative income." It turns out that the result obtained by Marhuenda and Ortuno-Ortin still holds. In other words, majority support for progressive taxes under absolute income concerns of the voters is not altered when relative income concerns are introduced. Mitra and Ok [63] prove that the principle of equal sacrifice also imply tax progressivity. In fact, they prove something stronger. They show that any marginal progressive tax is an equal sacrifice tax. However, all these approaches ignore incentive effects of taxation, incorporating which would be a natural but daunting next step.

There are other theoretical models supporting convexity as well. Roemer [72] looks at the problem, not from the voters' side, but from the side of the politicians. He comes up with a model and a new solution concept where the political parties would propose only progressive taxes in equilibrium (and hence the winner must be a progressive one). His model draws on factional conflicts within

each party. Each party consists of three factions - reformists, militants and opprtunists. However voters supply labor inelastically (no incentive effect of taxation). His policy space is that of quadratic (tax) functions. He then proves that in equilibrium (Strong Party Unanimity Nash Equilibrium), if most voters have income less than mean income, then both parties propose progressive tax policies. The obvious question is whether it is true if we introduce labor-leisure choices on the part of the voters. Roemer says that the result will hold if elasticity of labor supply with respect to wage for all incomes and all individuals is bounded above by some number and that number is very small. If not, he conjectures that progressive policies would not be necessarily proposed by either party.

Carbonell-Nicolau and Ok [17] have proved the existence of mixed-strategy equilibria and identified certain cases where marginal-rate progressive taxes are chosen almost surely by political parties. However, they too assume exogenous income distribution in the population (and therefore no incentive effect) and find that if the tax policy space is not artificially constrained, the support of at least one equilibrium cannot be contained within the set of marginal-rate progressive taxes.

The other direction of study has been the shape of 'optimal taxation', progressive or not, but with incentive effects. However it seems hard to model progressive taxes as being optimal. The pioneering attempt in this direction was that of Mirrlees [62]. His model has a social planner choosing a tax schedule that maximized social welfare subject to some budget constraint and incentive constraint of the voters (he does have labor-leisure choice). He arrives at the "striking" result of "closeness to linearity" of the optimal tax schedules. They are "not exactly linear" (but more concave) and "the maximum marginal rate occurs at a rather low income level, and falls steadily thereafter." However, he concludes from many examples, that the shape of the optimal schedule will depend on distribution of skills in the population and on the income-leisure preferences postulated.

Some authors, like Diamond [31] and Dahan and Strawczynski [27], have assumed away income effects (by assuming quasi-linear preferences), and have deduced that with Pareto and lognormal distributions marginal tax rates rise with income at high levels of income (conforming to empirical evidence). However if we do have income effects, such unambiguous characterization of optimal schedules is not possible analytically. Tuomala [85], [86] uses numerical simulations to conclude that with lognormal distribution (like Mirrlees) but with utility function quadratic in consumption, optimal tax rates imply rising marginal taxes. So with incentive effects, it seems like there could be convex schedules in equilibrium, in a benevolent social planner environment. So the question is, whether there can be such equilibrium schedules in a political environment.

However convexity does have a strong normative appeal on grounds of social justice. It is perhaps not aesthetically pleasing for any society to have the richest paying the lowest marginal taxes. So it doesn't seem to be an unreasonable assumption to begin by considering convex tax schedules only. The rest of the paper is organized as follows: Setion 2 introduces the model, the notation and the definitions; Section 3 talks about the benchmark case of the model with exogenous income and proves the non-existence of pure strategy equilibrium in that case; Section 4 talks about the model with endogenous income; Section 5 summarizes the main findings and concludes.

4.2 Notation and Definitions

The population consists of three types of people - one that has productivity a_1, the second group has productivity a_2, and the third group has productivity a_3. Let $a_1 > a_2 > a_3$ and $a_1 < 1$ and $a_3 > 0$. Let the size (fraction) of the population with productivity a_i be π_i, $i = 1, 2, 3$ such that $\pi_i < 1/2$ for all i and the sum of

the sizes of any two groups is greater than 1/2. We could variously perceive of the groups being the high, middle and low income groups. Alternatively they can constitute the public sector, the private sector with the third group consisting of retired/unemployed people (the least productive group), say.

Political Parties: There are two political parties that compete with each other for votes by proposing tax schedules for such a population. Whoever gets a majority of votes (that is, the votes of at least two groups in the population), wins with probability 1 and implements the announced platform (assuming election with commitment). Payoff of the candidate is not explicitly modeled here as it is enough for our purposes to assume that payoff is higher if he wins the election.

Voters: Each voter behaves as a utility maximizer under each of the two tax announcements (one from each candidate) and votes for the schedule that gives him higher utility. Let the utility function be $u(a_i, L, t(.)) = \ln(a_i L - t(a_i L)) + \ln(1 - L)$ where L is the amount of labor. $a_i L$ is the total income earned if an a_i productivity man works for L hours. The maximum possible amount of labor is normalized to be 1 so that $(1 - L)$ is the amount of leisure. We assume that candidates cannot observe effort levels but only earned incomes, so that tax announcements have to be based on observable earned incomes. $t(x)$ stands for the tax to be paid if earned income is x, under tax schedule $t(.)$. Hence $(a_i L - t(a_i L))$ is the amount of after-tax or disposable income. Assuming no savings etc, disposable income equals consumption, C. Hence utility positively depends on consumption, C and leisure, $(1 - L)$. Also the utility function has all the desirable properties like strict quasi-concavity and twice continuous differentiability.

Given tax announcement $t(.)$, an a_i-individual maximizes $u(.)$by choosing L. Since a_i is same for all people within a group, they all end up choosing the same optimal L. Different groups however choose different optimal L's.

Tax Schedule: A tax scedule, in general, is a continuous and increasing function $t : \mathbb{R}_+ \to \mathbb{R}$ such that (i) $0 \leq t(x) < x$ and (ii) $x - t(x)$ is increasing for all $x \in [0, 1]$ (where the second requirement ensures that the relative ranking of the individuals remains the same after taxes). Therefore $t(0) = 0$ and all tax schedules must begin from the origin. Diagrammatically, the first condition means that the tax schedule lies strictly below the 45^o line in the income-tax plane and the second condition means that its slope never exceeds 1. Let $\mathcal{T}_{conv}$ be the set of all tax schedules that satisfy the above conditions and are convex. Let $t(x) = y$. Let tax revenue paid by an individual of group i be y_i (and this is same for all individuals in that group). Tax revenue must also be at least R so that the revenue constraint becomes

$$R = \pi_1 y_1 + \pi_2 y_2 + \pi_3 y_3 \tag{4.1}$$

Let $\mathcal{T}_{conv}(R)$ be the set of all tax schedules in $\mathcal{T}_{conv}$ that collect exactly R.

The Game: Our game G is then as follows: the two candidates, 1 and 2, first announce two tax schedules, both members of $\mathcal{T}_{conv}(R)$, as functions of observable levels of earned incomes. An individual votes for that schedule that gives him higher utility. If his utility remains the same under both regimes, he votes for each party with probability 1/2. The party that gets more than half the votes, wins. If each gets exactly 1/2 the votes, each party wins with equal probability.

Equilibrium: The concept of equilibrium that we use is one of Nash Equilibrium. That is, (t_1^*, t_2^*) is a Nash Equilibrium if, given candidate 1 announces t_1^*, candidate 2 best responds by announcing t_2^* and vice versa. Hence, heuristically speaking, in such an equilibrium, if there is one, candidate 1 can always announce the Condorcet winner (if there is one) in the set of all feasible schedules and beat (or at least tie with) all announcements of candidate 2. Candidate 2 thinks likewise and proposes the Condorcet winner as well. Hence the (common) equilibrium schedule must be the Condorcet winner in the set of all feasible tax schedules.

4.3 A Benchmark: Exogenous Income

To facilitate comparison with the case of endogenous income, let earned income be exogenous in this case. That is, let people of all productivity choose the same amount of labor, say, $\bar{L}$ (which we consider to be 1 even though it makes the second term of our logarithmic utility function minus infinity, since we can just think of the utility containing the consumption term only). Since utility is assumed to depend on consumption and leisure, here utilities will differ only to the extent that their consumptions differ since leisure will be the same for people of all productivity types. Also, with tax announcements not influencing labor supply decisions, the only way taxes affect utility is by affecting consumption (negatively). Hence given any productivity type, lower the tax to be paid, higher will be consumption and higher will be the utility.

In our framework, since there are just three possible levels of earned income (three possible levels of productivity with labor equal to $\bar{L} = 1$), it is enough for the tax schedule to be just a three point specification - tax $t(a_i\bar{L}) = t(a_i)$ to be paid if income is $a_i\bar{L} = a_i$, $i = 1, 2, 3$. However, we must also keep in mind that three chosen points must preserve marginal progressiveness (convexity).

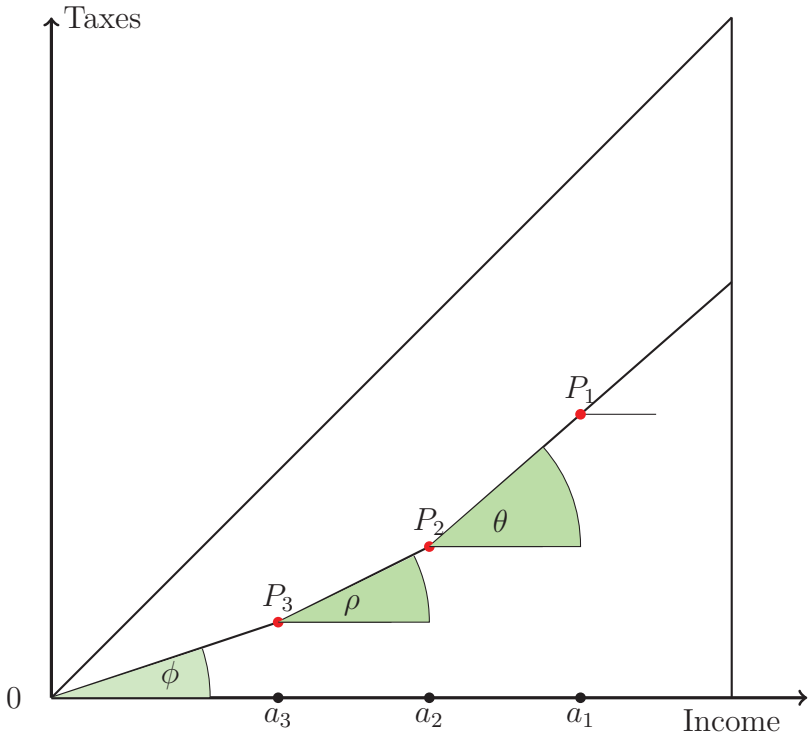


Figure 4.1. Graphical Representation

Refer to Figure 4.1. So let our generic tax schedule be given by $t(P_1, P_2, P_3)$ where $P_3 = (a_3, a_3 \tan\phi)$, $P_2 = (a_2, \tan\rho(a_2 - a_3) + a_3 \tan\phi)$ and $P_1 = (a_1, \tan\theta(a_1 - a_2) + a_2 \tan\rho)$ such that $\phi \leq \rho \leq \theta < \pi/4$ (for convexity). Then for revenue constraint to hold

$$R = \pi_1(\tan\theta(a_1 - a_2) + a_2 \tan\rho) + \pi_2(\tan\rho(a_2 - a_3) + a_3 \tan\phi) + \pi_3 a_3 \tan\phi$$

Hence for any such announcement t of one of the candidates, consider the opponent trying to come up with an alternative that makes any two groups better off, say 2 and 3. Then y_2 and y_3 must fall and y_1 increase appropriately such that R is still collected. For small changes so that first order approximations are accurate enough, we must have

$$\begin{aligned} dR = (\pi_3 a_3 + \pi_2 a_3)\sec^2\phi\, d\phi + (\pi_2(a_2 - a_3) + \pi_1 a_2)\sec^2\rho\, d\rho \\ + \pi_1(a_1 - a_2)\sec^2\theta\, d\theta \end{aligned}$$

For our purposes, we can pick $d\phi < 0$, $d\rho = 0$ (this serves to makes 2 and 3 better off) and choose θ so that

$$d\theta = -\frac{(\pi_3 a_3 + \pi_2 a_3)\sec^2\phi}{\pi_1(a_1 - a_2)\sec^2\theta} d\phi$$

which turns out to be positive. Note that for $\theta < \pi/4$ (which is true by the assumption that the marginal tax rate is always less than 100%), we can choose $d\phi$ small enough so that $\theta + d\theta < \pi/4$. The resulting schedule makes 2 and 3 better off and 1 worse off and therefore wins in majority voting over the previous schedule.

If $\phi < \rho$, then we could similarly construct alternatives that make 1 and 3 better off (in fact we're going to do it in the endogenous income case), or 1 and 2 better off. However if $\phi = \rho$, there will be only two possibilities of making any two groups better off without violating convexity and still collecting R, making 2 and 3 better off or 1 and 3 better off (but not 1 and 2). Hence, we get the (already well-known) result stated in the following proposition:

Proposition 3 *There is no pure strategy equilibrium of the game G when income is exogenous.*

4.4 The Model with Endogenous Income

This is otherwise similar to the previous model except that we let individual voter now choose between labor and leisure for any given tax announcement. Such decisions, we assume, are not observable by the candidates (and hence the question of incentive-compatible tax-schedules). The individuals make their labor choices (through maximizing utility) under the two tax regimes announced by the candidates and hence find out what their utilities are going to be if one or the other schedule is actually implemented. (With utility function being the same

for everyone, and productivity being the same within each group, each member of the same group ends up making the same labor choice.)

Now all basic results regarding existence of solution to the maximization problem of the voter, the maximum revenue that can be collected through convex schedules, etc are relegated to the appendix (please see appendix .5 Taxation-A).

Our objective is to study existence and properties of equilibrium in such an environment (where an equilibrium would exist if there is a Condorcet winner in the feasible set of tax schedules). We try to zero-in on the equilibrium schedule with the following broad steps:

Step 1: We first show that we can consider a subset of the set of all feasible schedules and call these the *3-point* tax schedules. The first step consists in characterizing this subset and is summarized in *Lemmas 4, 6, 11* and *12*.

Step 2: We then show even in this subset we can restrict attention to a further subset that satisfy *'Property α'* (since all non-α schedules are Pareto-dominated). (*Proposition 13*)

Step 3: We then show that if a tax schedule does not satisfy *Property β*, then there is a β-schedule that both groups 1 and 2 prefer. Hence any equilibrium tax schedule must satisfy β. (*Proposition 14*)

Step 4: Among all β-satisfying schedules, there is one, say $\hat{t}$, which maximizes the utilities of groups 2 and 3. (*Proposition 16*) Hence if equilibrium exists, it must be $\hat{t}$.

Step 5: We look at $\hat{t}$ more closely and try to characterize situations when it is indeed an equilibrium (and a linear one too) and when it is not. (*Theorem 17* and *Theorem 19*)

Step 6: Finally we illustrate for what parameter values a strictly convex equilibrium $\hat{t}$ exists. (*Theorem 21*)

STEP 1:

Definition 4 *We call a tax schedule a 3-point tax schedule if it belongs to* $\mathcal{T}_{conv}$ *and has* 3 *kinks, say at the points* P_1, P_2 *and* P_3 *respectively. (Refer to figure 4.2.)*

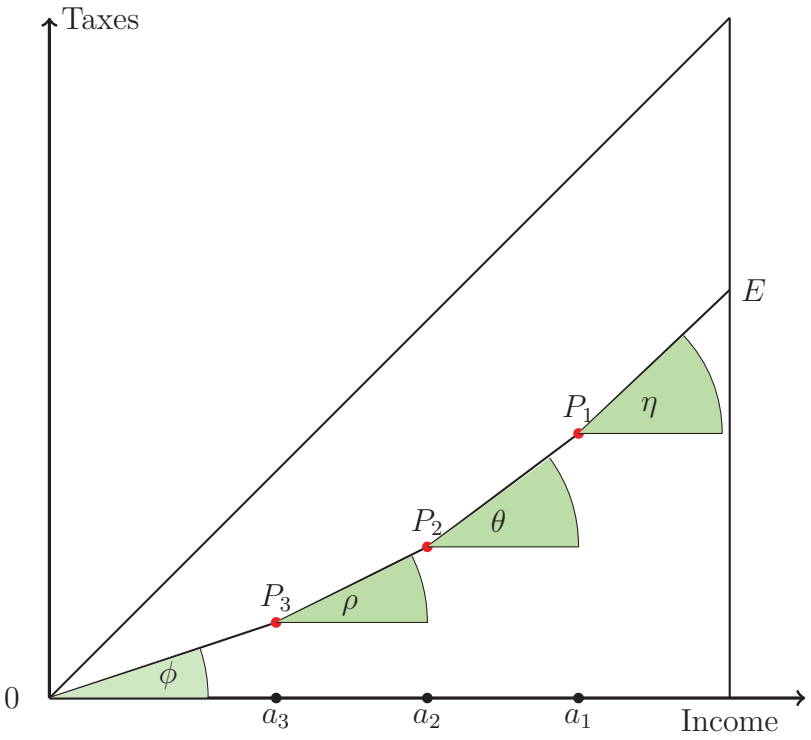


Figure 4.2: 3−point Tax Schedule

Definition 5 *Two tax schedules are said to be equivalent if both of them collect the same revenue and yield the same levels of utilities to all the groups.*

Lemma 6 *Let* t *and* $t^* \in \mathcal{T}_{conv}$ *such that* $t^*(x) \geq t(x)\ \forall x$. *Suppose a particular productivity group* α, *chooses income* p *under tax schedule* t *and suppose* $t^*(p) = t(p)$. *Then* α *chooses the same income* p *under* t^*.

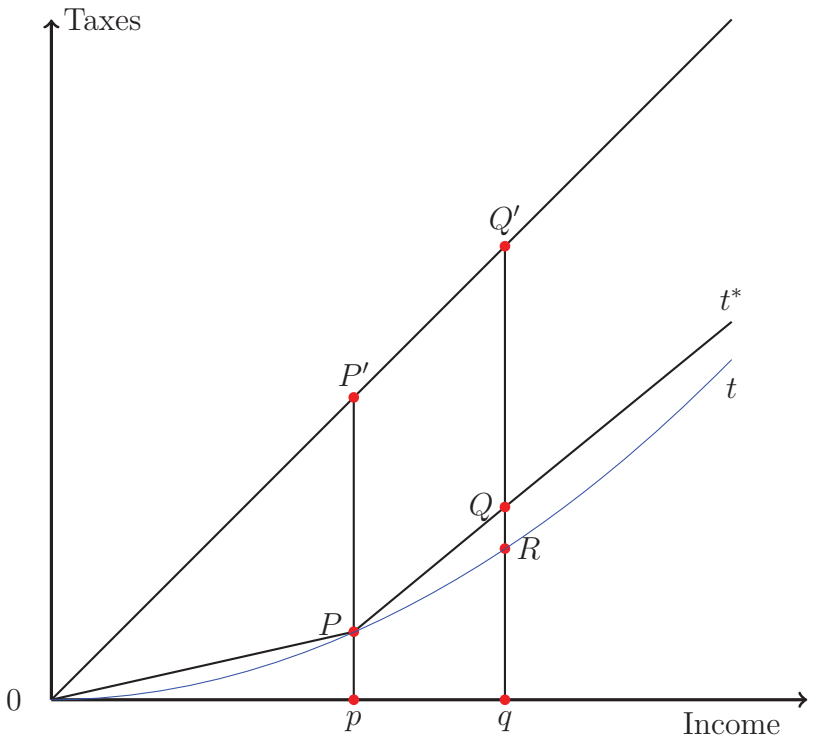


Figure 4.3. Proof of Lemma 4

Proof. We provide a geometric proof. Refer to figure 4.3. Suppose, by way of arriving at a contradiction, the optimal labor choice for α under t^* be such that the earned income is q which corresponds to point Q on t^*. Let the line segment qQ intersect t at R. Now if Q is chosen under t^* when P was available, by revealed preference, $U_Q > U_P$ (where U_X stands for the utility at the point X). Since R is any other point on t and P is the optimal, we have $U_P > U_R$. By transitivity, $U_Q > U_R$.

Now utility is increasing in leisure and consumption. But $\frac{\alpha L}{\alpha}$ are the same at the points Q and R given by the ratio of Oq to α, so leisure $1 - L = 1 - \frac{\alpha L}{\alpha}$ are same at the two points. Now since $t^*(q) \geq t(q)$, we get $Q'R$ is bigger than $Q'Q$ i.e. consumption is higher at R than at Q. So $U_R \geq U_Q$ - a contradiction. ■

Lemma 4 tells us that if tax for my labor choice is unchanged and that for any other labor choices have increased, I keep my labor choice unchanged.

Corollary 7 *For any $t \in \mathcal{T}_{conv}$ there exists a 3-point curve $t^* \in \mathcal{T}_{conv}$ that is equivalent to t.*

Proof. Suppose group a_i chooses P_i, $i = 1, 2, 3$ under t. Form a 3-point curve such that the kinks are at P_1, P_2 and P_3 and call it t^*. Then it follows from the above lemma that t and t^* are equivalent. ■

Lemma 8 *A higher productivity group is better off at a higher kink than at a lower one.*

Proof. Suppose not. Suppose group 2 is at P_1 so that it now chooses (x_1, y_1), and group 1 is at P_2 who now chooses (x_2, y_2). Now extend $P_1 - P_2$ till it reaches the horizontal axis (refer to Figure 4.4). Now suppose the tax schedule would be OPR.

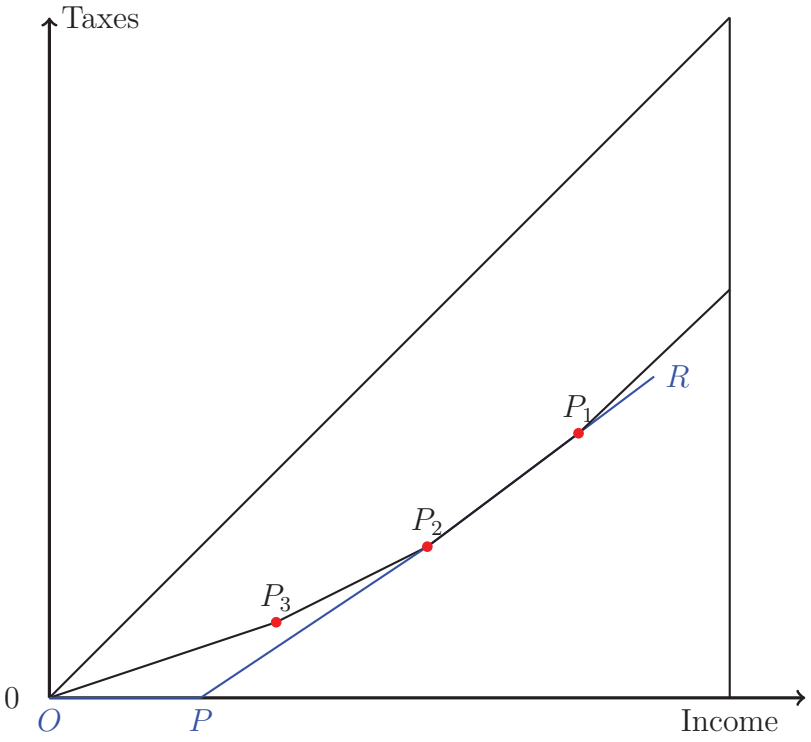


Figure 4.4. Proof of Lemma 6

We can check that the optimal for productivity a_i for this schedule would be as

follows:

$$\text{If} \quad a_i > 2P + \frac{P\tan\psi}{1-\tan\psi}, \quad \text{then} \quad L = \frac{1}{2} - \frac{P\tan\psi}{2a_i(1-\tan\psi)}$$
$$\text{If} \quad 2P + \frac{P\tan\psi}{1-\tan\psi} \geq a_i \geq 2P, \quad \text{then} \quad L = \frac{P}{a_i}$$
$$\text{If} \quad 2P > a_i, \quad \text{then} \quad L = \frac{1}{2}$$

What is important is that we see that earned income $x = a_i L_i$ positively depends on a_i.

Now given the P_1P_2 segment, if 1 is at P_2, then 1's optimal point (when given OPR) must be below P_2 (since it is not in P_2P_1 else he would choose that given it is available, and it is not above P_1 since then he would be better off choosing P_1 than P_2). Now given 1's choice, 2's must be below 1's (given optimal labor choices of the two groups). Hence 2 would also choose P_2 when given P_2P_1 which is a contradiction to 2 choosing P_1. ■

Definition 9 *A 3-point tax schedule is called permissible if the kinks correspond to the optimal points of the three groups respectively. Call the set of such tax schedules $\mathcal{T}^*_{conv}$. Call it $\mathcal{T}^*_{conv}(R)$ if the schedules also collect R.*

Remark 10 *Not all 3-point tax schedules are permissible because the kinks may not be the optimal points always.*

Remark 11 *Corollary 5 tells us that instead of studying infinite dimensional space of tax schedules $\mathcal{T}_{conv}$, it is enough to consider the five dimensional space $\mathcal{T}^*_{conv}$ (3-point tax schedules are defined by six coordinates with y_1, y_2 and y_3 related through the revenue constraint).*

Remark 12 *By Lemma 6, higher productivity groups must be at higher kinks. So that if three groups had to be at three kinks, it must be the case that 1 is at the highest kink (rightmost), 3 is at the lowest (leftmost) and 2 is in between the*

two. Hence higher the productivity, higher must be earned income (and higher must be taxes given we have assumed increasing taxes). (A similar result holds for differentiable tax schedules, see Claim 24 in appendix Taxation-A).

A complete characterization of such schedules is given next.

Criteria for Permissibility

Our next task is to find criteria for a 3-point tax schedule to be permissible in terms of the coordinates of the kinks. We denote the kinks by P_1, P_2, P_3 from right to left (like in figure 4.2) with coordinates (x_1, y_1), (x_2, y_2) and (x_3, y_3) respectively where X-axis measures earned income and Y-axis measures taxes. It begins at the origin $O = (0, 0)$ and ends at $E = (1, e)$.

Lemma 13 *P_i $\forall i$, is optimal for any 3-point tax schedule if it is optimal in the two adjacent segments containing P_i $\forall i$.*

Proof. Suppose not. Consider group 1. Suppose it is optimal for 1 to choose P_1 in the two adjacent segments containing P_1 but not optimal when the whole tax schedule is available. Suppose then a point on the P_3P_2 is chosen. Since income is increasing in productivity, it means that group 2 chooses somewhere below, but he cannot choose on the P_3P_2 segment since P_2 is optimal for 2 on the P_3P_2-P_2P_1 segments and P_2 is higher than what 1 has chosen. Hence let 2 choose on the OP_3 segment and 3 settles at a even lower point on OP_3. This means P_3 is not optimal for 3 in the OP_3-P_3P_2 segments - a contradiction to hypothesis. Similarly for other cases and groups. ■

Lemma 14 *A 3-point tax schedule is permissible if and only if all the following*

conditions are satisfied:

$$\frac{a_1}{2} + \frac{ex_1 - y_1}{2(e + x_1 - y_1 - 1)} \leq x_1 \leq \frac{a_1}{2} + \frac{y_1x_2 - x_1y_2}{2(y_1 + x_2 - y_2 - x_1)} \tag{4.2}$$

$$\frac{a_2}{2} + \frac{y_1x_2 - x_1y_2}{2(y_1 + x_2 - y_2 - x_1)} \leq x_2 \leq \frac{a_2}{2} + \frac{y_2x_3 - x_2y_3}{2(y_2 + x_3 - y_3 - x_2)} \tag{4.3}$$

$$\frac{a_3}{2} + \frac{y_2x_3 - x_2y_3}{2(y_2 + x_3 - y_3 - x_2)} \leq x_3 \leq \frac{a_3}{2} \tag{4.4}$$

Proof. The FOCs for (unique) maximization for the different groups along with the optimality conditions in Lemma 11 give rise to these inequalities. Please see appendix .6 Taxation-B for a proof of the proposition in case of group 2. For groups 1 and 3 the arguments are similar. ■

It will perhaps be a little tidier to write the coordinates in terms of the relative angles between the segments. Let the OP_3 segment make angle ϕ with the horizontal axis, P_3P_2 segment make angle ρ with the horizontal axis, P_2P_1 segment make angle θ with the horizontal axis and P_1E segment make angle η with the horizontal axis. Convexity implies $0 < \phi \leq \rho \leq \theta \leq \eta < \pi/4$. Then the above conditions can be written as

$$\frac{a_1}{2} - \frac{\tan\eta - e}{2(1 - \tan\eta)} \leq x_1 \leq \frac{a_1}{2} - \frac{x_2\tan\theta - y_2}{2(1 - \tan\theta)} \tag{4.5}$$

$$\frac{a_2}{2} - \frac{x_1\tan\theta - y_1}{2(1 - \tan\theta)} \leq x_2 \leq \frac{a_2}{2} - x_3\frac{\tan\rho - \tan\phi}{2(1 - \tan\rho)} \tag{4.6}$$

$$\frac{a_3}{2} - \frac{x_2\tan\rho - y_2}{2(1 - \tan\rho)} \leq x_3 \leq \frac{a_3}{2} \tag{4.7}$$

Note that these conditions imply incentive compatibility for the groups so that no group would want to choose any other point on the curve by deviating from its own.

STEP 2:

Property α: A tax schedule $t \in \mathcal{T}^*_{conv}$ satisfies α-property if

$$x_1 = \frac{a_1}{2} - \frac{x_2 \tan\theta - y_2}{2(1 - \tan\theta)}$$

Satisfaction of this property means that utility of group 1 is maximized when it is given line P_2P_1.

Proposition 15 *If a tax schedule $t \in \mathcal{T}^*_{conv}(R)$ does not satisfy α-property, then we can find another tax schedule in $\mathcal{T}^*_{conv}(R)$ (satisfying α-property) that does strictly better i.e. makes all the groups better off.*

Proof. Pick a schedule t' in $\mathcal{T}^*_{conv}(R)$ that does not satisfy α-property. Then it means

$$x_1 < \frac{a_1}{2} - \frac{x_2 \tan\theta - y_2}{2(1 - \tan\theta)}$$

Suppose now a tax schedule is proposed that is exactly like t' except that the kink at P_1 corresponds to $\frac{a_1}{2} - \frac{x_2 \tan\theta - y_2}{2(1-\tan\theta)}$. Then 1 pays more taxes and 2 and 3 pay the same (and 1 is strictly better off). Hence revenue collected is more than R. Hence we can define a family of tax schedules by slightly reducing ϕ (keeping the other angles same) and $x_1 = \frac{a_1}{2} - \frac{x_2 \tan\theta - y_2}{2(1-\tan\theta)}$. Note that 1, 2 and 3 are better off in each curve of this family compared to t'. Due to continuity of revenue, there will be a curve that gives exactly R and makes all the three groups better off. ■

Therefore, from Proposition 13 we know that all the tax schedules in $\mathcal{T}^*_{conv}(R)$ that do not satisfy α-property are strictly dominated (in the sense of gaining majority support) by some other in $\mathcal{T}^*_{conv}(R)$ and hence can be dropped for our immediate purpose. Hence, let $\mathcal{T}^{*\alpha}_{conv}(R)$ denote schedules where α-property is satisfied in addition. This again means that for all practical purposes, the fourth angle, η, can easily be dropped and the third angle can be continued till $x = 1$. So our tax schedule will look like this:

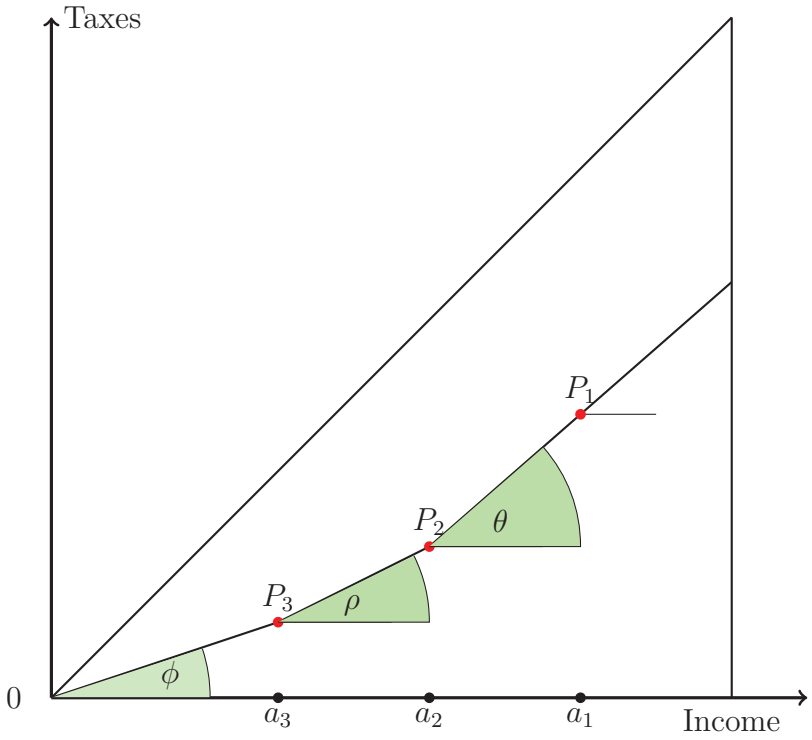


Figure 4.5. Reduced 3-point Tax Schedule

STEP 3:

Property β: A tax schedule $t \in \mathcal{T}^*_{conv}$ satisfies β-property if the point P_3 lies on the line segment joining the origin and P_2. i.e. if $\phi = \rho$.

Proposition 16 *For any $t \in \mathcal{T}^{*\alpha}_{conv}(R)$ not satisfying property β, there exists a $t^* \in \mathcal{T}^{*\alpha}_{conv}(R)$ satisfying property β such that $t^* \succeq_{MR} t$ (i.e. t^* beats t in majority voting) by attracting groups 1 and 2.*

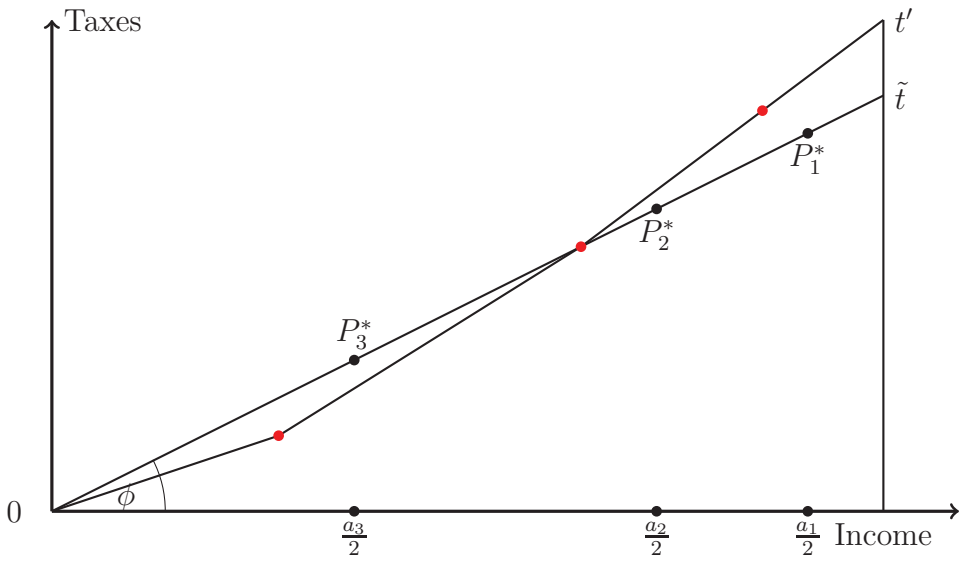


Figure 4.6. Proof of Proposition 14

Proof. Pick any $t' \in \mathcal{T}^{*\alpha}_{conv}(R)$ such that $t' = t(P_1, P_2, P_3)$. Construct $\tilde{t}$ such that it is a straight line passing through the origin and the point P_2 (as a result of which, with abuse of notation, it makes angle ϕ again, with the origin). Then, putting $\phi = \rho = \theta$ yields that utility maximizing income choices will be $x_i = a_i/2$ under $\tilde{t}$. (In fact optimal income choice under any straight line tax schedule for any group a is $a/2$.) Let the corresponding points on $\tilde{t}$ be called P_i^*. Note that under $\tilde{t}$, groups 2 and 3 pay higher taxes while 1 pays lower taxes so that the net effect on total tax collected is ambiguous.

Case 1: $R_{\tilde{t}} > R$ (Revenue collected under $\tilde{t} > R$)

$$\begin{aligned} R_{\tilde{t}} &= \sum_{i=1}^{3} \pi_i (a_i/2) \tan\phi \\ &= (\tan\phi/2)\mu \\ &> R \end{aligned}$$

where $\mu = \sum_{i=1}^{3} \pi_i a_i$.

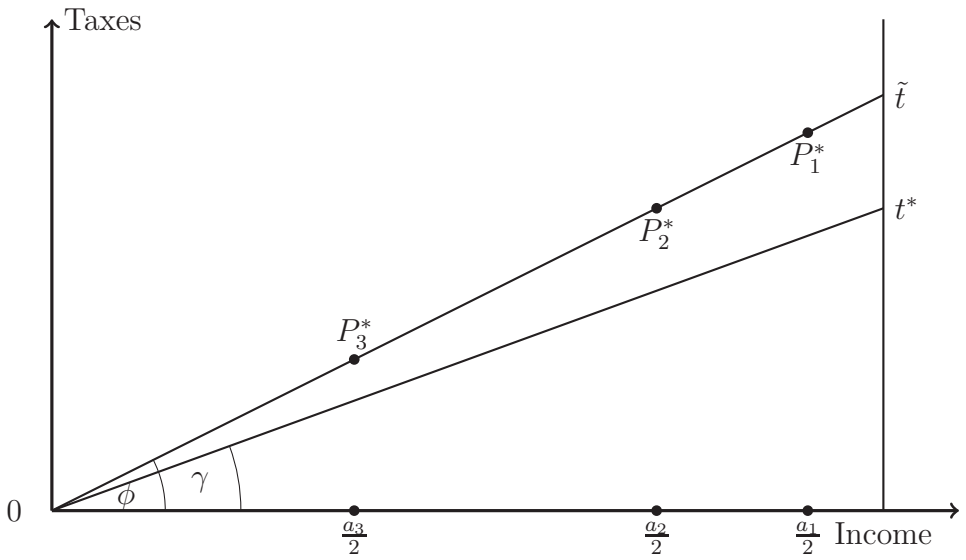


Figure 4.7. Case 1

Now construct t^*, a straight line like $\tilde{t}$ but now making angle γ with the origin such that γ satisfies $R = (\tan\gamma/2)\mu$ (hence pick $\gamma = \tan^{-1}(2R/\mu)$). Hence $t^* \in \mathcal{T}^*_{conv}(R)$. Now groups 1 and 2 are better off under t^* than under t' and hence $t^* \succeq_{MR} t'$. Note that in case $P_2 = P_2^*$, the whole argument still goes through. Also note that any straight line tax schedule (t^* in particular) satisfy the β property. Hence the proposition is true in this case.

Case 2: $R_{\tilde{t}} < R$ and $P_2 \neq P_2^*$

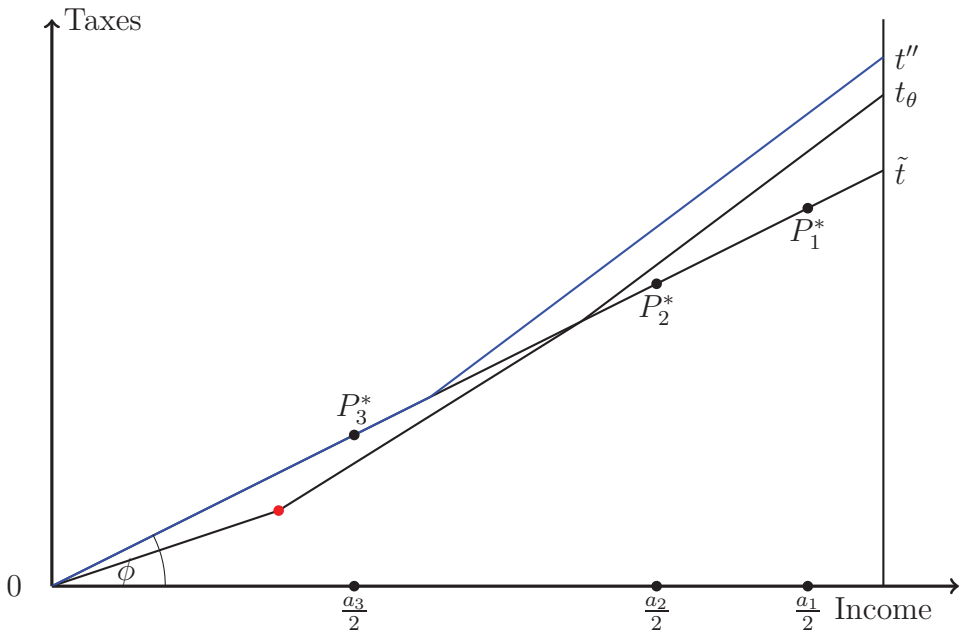


Figure 4.8. Case 2

Let t'' be the tax schedule that is obtained by taking t' and making it satisfy property β. Hence $t'' = t(P_1, P_2, P_3^*)$. Notice that t'' collects more than R (since group 3 pays higher, 2 and 1 pay the same). Let θ_0 be the angle made by the segment P_2P_1 in the original tax schedule t' (which is same as in t''). We construct a family of tax schedules, t_θ, parameterized by $\theta \in [\phi_0, \theta_0]$. The tax schedules will be given by joining $O\hat{P}_2$ where $\hat{P}_2 = (\hat{x}_2, \hat{x}_2 \tan\phi_0)$ and thereafter making angle θ.

Equation of the tax schedule after $\hat{P}_2$ is given by

$$y = \tan\theta(x - \hat{x}_2) + \hat{x}_2 \tan\phi_0$$

Computations, similar to those done before, yield, that utility is maximum for group 2 at

$$x = \frac{a_2}{2} - \hat{x}_2 \frac{\tan\theta - \tan\phi_0}{2(1 - \tan\theta)}$$

For IC constraint of group 2 to hold, we need,

$$\hat{x}_2 \geq \frac{a_2}{2} - \hat{x}_2 \frac{\tan\theta - \tan\phi_0}{2(1 - \tan\theta)}$$

which implies

$$\hat{x}_2 \geq \frac{a_2(1-\tan\theta)}{2-\tan\theta-\tan\phi_0}$$

Note that the right hand side is a decreasing function of θ. For given θ we choose

$$\hat{x}_2 := \max\left\{\frac{a_2(1-\tan\theta)}{2-\tan\theta-\tan\phi_0}, x_2\right\}$$

Thus we get a family of tax schedules, t_θ, satisfying β property, parameterized by $\theta \in [\phi_0, \theta_0]$ such that $t_{\phi_0} = \tilde{t}$ and $t_{\theta_0} = t''$. Since $R_{t''} > R > R_{\tilde{t}}$ (in this case), there exists a θ^* such that $R_{t_{\theta^*}} = R$ (by continuity of revenue) and groups 2 and 1 are better off under $t^* := t_{\theta^*}$ compared to t'.

Case 3: $R_{\tilde{t}} = R$ and $P_2 \neq P_2^*$

Choose t^* to be $\tilde{t}$. Then it satisfies α and β, collects R and groups 1 and 2 are better off under t^* than under t'.

Case 4: $R_{\tilde{t}} \leq R$ and $P_2 = P_2^*$

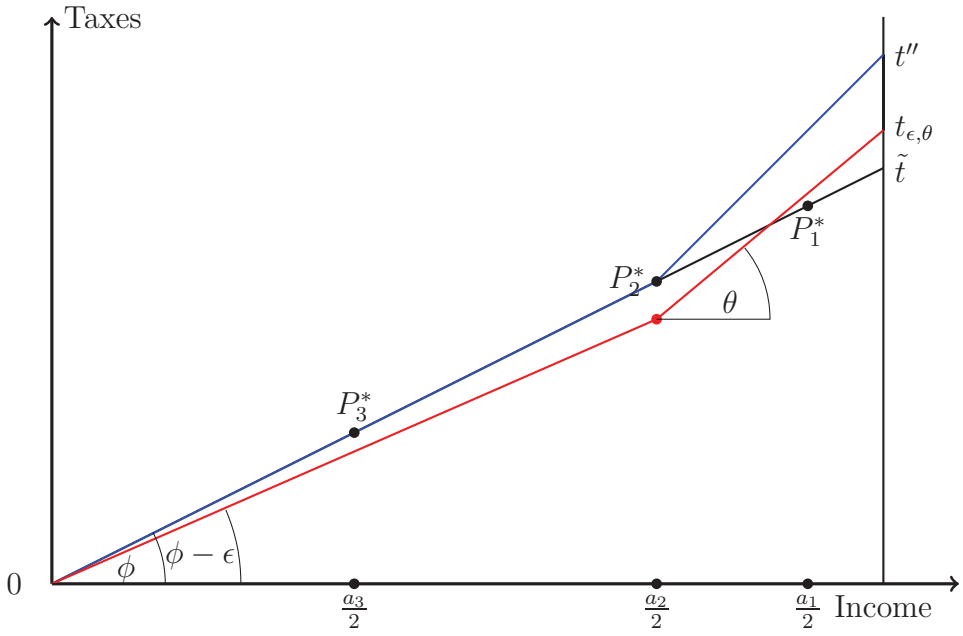


Figure 4.9. Case 4

We construct a family from t'' in the following way: we reduce the angle ϕ by ϵ and then extend it at angle θ (notice the abuse of notation) from $x = a_2/2$ (the x corresonding to P_2^*) onwards. Hence (by α condition) we have

$$\hat{x}_1 = \frac{a_1}{2} - \frac{a_2}{2}\frac{\tan\theta - \tan(\phi_0 - \epsilon)}{2(1-\tan\theta)}$$

Analyzing like before we can check that since $\phi_0 - \epsilon \leq \theta$, the incentive compatibility for group 2 is satisfied.

Therefore, we get a continuous family of tax schedules $t_{\epsilon,\theta}$ such that $t_{0,\theta_0} = t''$, $R_{t''} > R$, $t_{0,\phi_0} = \tilde{t}$, and $R_{\tilde{t}} \leq R$.

Consider the tax schedule $t_{\phi_0/2,\theta_0}$. If this collects revenue more than R then there exists a tax schedule in the one parameter family $\{t_{\phi_0/2,\theta} : \theta \in [\phi_0/2, \theta_0]\}$, which collects revenue R and groups 1 and 2 are better off. On the other hand if $t_{\phi_0/2,\theta_0}$ collects revenue less than R then there exists a tax schedule in the one parameter family $\{t_{\epsilon,\theta_0} : \epsilon \in [0, \phi_0/2]\}$, which collects revenue R and groups 1 and 2 are better off. ■

Remark 17 *Hence if there were to be an equilibrium, it must be one like t^* in Proposition 14 or else any candidate could choose it and defeat any other.*

STEP 4:

Proposition 18 *Out of all $t \in \mathcal{T}^{*\alpha}_{conv}(R)$ satisfying β, there exists a unique t, call it $\hat{t}$, that maximizes the utilities of groups 2 and 3.*

Proof. A typical curve in this set is parameterized by the point $P_2 = (x_2, y_2)$, since this gives ϕ and θ is then given by the revenue constraint. Assume $x_2 > a_3/2$. Then the points P_1 and P_3 (in terms of x_2 and y_2) will be

$$P_1 = \left(\tfrac{a_1}{2} - x_2\tfrac{\tan\theta-\tan\phi}{2(1-\tan\theta)}, \left[\tfrac{a_1}{2} - x_2\tfrac{\tan\theta-\tan\phi}{2(1-\tan\theta)} - x_2\right]\tan\theta + y_2\right)$$
$$P_3 = \left(\frac{a_3}{2}, \frac{a_3}{2}\tan\phi\right)$$

Note that x_1 (the first coordinate of P_1) is decreasing in θ. So as θ increases the point P_1 moves towards P_2. So an upper bound for θ will be defined by $x_2 = x_1$, call it $\bar{\theta}$. Also substituiting y_3 and y_1 in terms of x_2,and y_2 in the revenue constraint we get

$$R = \pi_1\Big(\Big[\frac{a_1}{2} - x_2\frac{\tan\theta - \tan\phi}{2(1-\tan\theta)} - x_2\Big]\tan\theta + y_2\Big) + \pi_2 y_2 + \pi_3\frac{a_3}{2}\tan\phi$$

This gives a quadratic equation in $\tan\theta$ which we can solve for the two values of θ, say θ_+ and θ_-. If both the values $\in [\phi, \bar{\theta}]$, we take the smaller value since that gives higher utility to group 1. If both the values lie outside $[\phi, \bar{\theta}]$, then (x_2, y_2) doesn't give a feasible tax schedule. If $\phi \leq \theta_+ \leq \bar{\theta}$ (and the other is not), it gives a closed set of (x_2, y_2) values and if $\phi \leq \theta_- \leq \bar{\theta}$, it also gives a closed set of values of (x_2, y_2). The union of these two sets of feasible values gives the closed set of all possible (x_2, y_2) values, say $\mathcal{C}$.

Let $\phi_o = \min\{\phi = \tan^{-1}(y_2/x_2) : (x_2, y_2) \in \mathcal{C}\}$. Note this exists since $\mathcal{C}$ is a closed set. And $\bar{x}_2 = \max\{x_2 : (x_2, y_2) \in \mathcal{C} \text{ and } \tan\phi_o = y_2/x_2\}$. Then we define our t^* to be that defined by $\tan\phi_o = y_2/\bar{x}_2$.

Claim 1: $\bar{x}_2 = a_2/2$

Proof: Suppose not i.e. $\bar{x}_2 < a_2/2$. (This also implies $\phi_o < \theta$ else if $\phi_o = \theta$, $x_2 = a_2/2$). Also suppose, for the time being $\theta < \bar{\theta}$. Then we claim that it is possible to increase x_2 ($\bar{x}_2$) by dx_2 thereby contradicting the definition of $\bar{x}_2$ and still collect R. Note that once ϕ_o and x_2 are given, y_2 is given, and y_1 becomes a function of θ and x_2. Also group 3's position does not change. So the change in

revenue when x_2 is changed is given by

$$\begin{aligned} dR &= \pi_2 \tan\phi_o dx_2 + \pi_1 dy_1 \\ &= \pi_2 \tan\phi_o dx_2 + \pi_1 \Big(\frac{\partial y_1}{\partial \theta} d\theta + \frac{\partial y_1}{\partial x_2} dx_2 \Big) \\ &= \Big(\pi_2 \tan\phi_o + \pi_1 \frac{\partial y_1}{\partial x_2} \Big) dx_2 + \pi_1 \frac{\partial y_1}{\partial \theta} d\theta \end{aligned}$$

For dR to be 0, pick

$$d\theta = - \frac{\Big(\pi_2 \tan\phi_o + \pi_1 \frac{\partial y_1}{\partial x_2} \Big) dx_2}{\pi_1 \frac{\partial y_1}{\partial \theta}}$$

which can be done whenever $\frac{\partial y_1}{\partial \theta} \neq 0$ (and since $\theta < \bar{\theta}$).

Now suppose $\bar{x}_2 < a_2/2$ but $\theta = \bar{\theta}$. Then P_1, P_2 are the same points and $x_1 = x_2 = \bar{x}_2$. Then if you give the straight line tax schedule with angle ϕ_o at the origin, it collects more than R (since 2's optimal point is $a_2/2$ which yields taxes higher than at $\bar{x}_2$ and group 1 gives more). But then ϕ_o is not the minimum angle, a contradiction to the definition of ϕ_o.

Also we've proved, $\bar{x}_2$ cannot be $< a_2/2$ when $\theta < \bar{\theta}$. Hence $\bar{x}_2 = a_2/2$. So tax schedule looks like $t(P_1, P_2, P_3)$ with $x_3 = a_3/2$, $x_2 = a_2/2$. This proves Claim 1.

Claim 2: Utility is maximized for group 3

Proof: For group 3 to get higher utility, they must pay lower taxes. This can happen only if we lower ϕ_o, but since it's the minimum we can't get any (x_2, y_2) in C consistent with it.

Claim 3: Utility is maximized for group 2

Proof: If not then there is a tax schedule (given $\mathcal{C}$) that gives higher utility to group 2. This means it is more towards the utility maximization point for group 2, P_2. But this means an angle lower than ϕ_o, a contradiction.

This proves the proposition. ■

Hence our candidate for equilibrium would be $\hat{t}$.

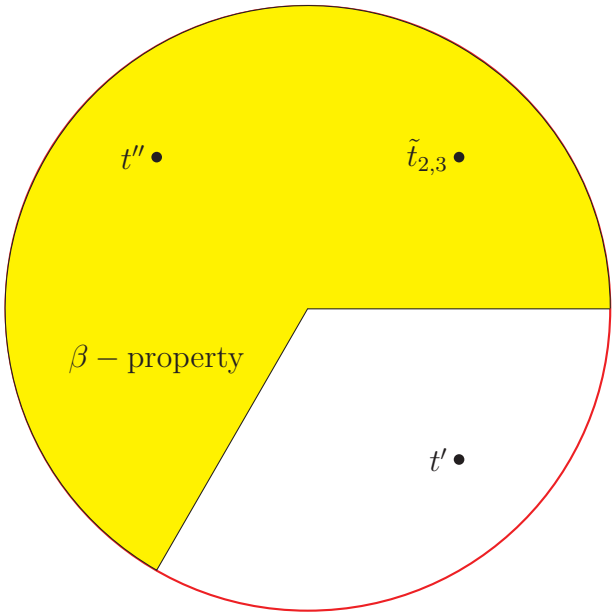


Figure 4.10. Who will Win?

So, as shown in figure 4.10, for any tax schedule t' in $\mathcal{T}^*_{conv}(R)$, one can choose some t'' (that satisfies α and β properties) and beat it by attracting groups 1 and 2. And any such schedule can be beaten by $\hat{t}$ by attracting groups 2 and 3.

STEP 5:

As θ changes, y_1 (the taxes paid by group 1) changes. How? $\frac{\partial y_1}{\partial \theta}$ yields a quadratic equation in $\tan\theta$ which vanishes for a particular value of θ. This means that as θ changes, y_1 first increases and then falls. So for every ϕ, we can find a θ that gives revenue R but this θ may not correspond to y_1 being maximal. However we claim that, for the minimal angle ϕ_o, the corresponding θ must be that which maximizes y_1 (and gives R). Let the θ corresponding to ϕ_o be θ_o.

Claim A: Either $\frac{\partial y_1}{\partial \theta}|_{\theta_o} = 0$ or $\theta_o = \phi_o$

Proof: Given $x_2 = a_2/2$, R is a function of ϕ and θ i.e. $R = R(\phi, \theta)$. Therefore, $dR = \frac{\partial R}{\partial \theta} d\theta + \frac{\partial R}{\partial \phi} d\phi$. Also only y_1 depends on θ (y_2 and y_3 don't). So $\frac{\partial R}{\partial \theta} = \pi_1 \frac{\partial y_1}{\partial \theta}$. So as long as $\frac{\partial y_1}{\partial \theta}$ is $\neq 0$, $\frac{\partial R}{\partial \theta}$ is not zero. Hence for $dR = 0$, we can find

$$d\theta = -\frac{\frac{\partial R}{\partial \phi} d\phi}{\frac{\partial R}{\partial \theta}}$$

Hence we can choose $d\phi$ negative and sufficiently small such that $\theta + d\theta \geq \phi + d\phi$ still holds (convexity is preserved) and revenue R is still obtained. But this contradicts the definition of ϕ_o and such a ϕ cannot be ϕ_o. Hence $\frac{\partial y_1}{\partial \theta}$ must be zero at θ_o corresponding to ϕ_o.

In all cases where the maximum y_1 occurs at an angle $\theta \leq \phi_o$, the θ has to chosen to be ϕ_o, so that $\theta_o = \phi_o$.

Solving for the angles

We can now solve for the angle θ that maximizes y_1 for every given ϕ. Setting $\frac{\partial y_1}{\partial \theta} = 0$, we get

$$\tan\theta = 1 - \sqrt{\frac{a_2(1 - \tan\phi)}{2a_1 - a_2}}$$

Note that if the schedule is not a straight line, θ_o exactly solves

$$\tan\theta_o = 1 - \sqrt{\frac{a_2(1 - \tan\phi_o)}{2a_1 - a_2}}$$

($\theta_o = \phi_o$ if the schedule is a straight line). By substituting for $\tan\theta$ in the revenue constraint we can find out what ϕ_o should be. It turns out to be the solution of the following equation

$$\sqrt{2a_1 - a_2} - \sqrt{a_2(1 - \tan\phi)} = 2\sqrt{\frac{R - A\tan\phi}{\pi_1}}$$

where

$$A = \frac{\pi_1 a_2 + \pi_3 a_3 + \pi_2 a_2}{2}$$

Straight line

We see that if $R = \frac{a_1 - a_2}{2a_1 - a_2}\mu$ then the resulting ϕ_0 is such that the corresponding maximizing θ_0 is exactly equal to ϕ_0.

Theorem 19 *For all $R \geq \frac{a_1 - a_2}{2a_1 - a_2}\mu$, $\theta_o = \phi_o$ (i.e. $\hat{t}$ is a straight line) and it is an equilibrium.*

Proof. : If $R = \frac{a_1 - a_2}{2a_1 - a_2}\mu$ then we know that the tax schedule will be a straight line. So for all R greater than that, it would require higher y_1 (or lower θ) but any lower θ would imply a concave curve and violate convexity. It will be an equilibrium since to do better than that either 2 and 3 have to be attracted by lowering taxes (i.e. by lowering ϕ) but that violates the minimality of angle ϕ_0 or 1 and 2 have to be attracted which means all three groups pay less (ϕ also has to be lowered) and R cannot be collected. Notice that with a straight line tax schedule it is not possible to maintain convexity and attract groups 1 and 3. ■

Remark 20 *The above proposition completely settles the case when $R \geq \frac{a_1 - a_2}{2a_1 - a_2}\mu$. That is, for revenue large enough equilibrium exists and it is a straight line tax schedule out of all convex schedules.*

Therefore hereafter we concentrate only on the case of $R < \frac{a_1 - a_2}{2a_1 - a_2}\mu$.

Property of $\hat{t}$

$\hat{t}$ maximizes 2's utility globally (in $\mathcal{T}^*_{conv}(R)$). This is because among all tax schedules in $\mathcal{T}^*_{conv}(R)$, those satisfying β give higher utilities to 1 and 2 (Proposition 14) and out of all schedules in $\mathcal{T}^*_{conv}(R)$ satisfying β, $\hat{t}$ gives maximum utilities to 2 and 3 (Proposition 16). Hence $\hat{t}$ gives maximum utility to 2 out of all schedules in $\mathcal{T}^*_{conv}(R)$.

Now consider the case where $\hat{t}$ is not straight. Then for all $t \in \mathcal{T}^*_{conv}(R)$ satisfying β (and α) , we know that a candidate can always announce $\hat{t}$ and win under majority rule. For there not to be an equilibrium, it must then be true that given $\hat{t}$, a candidate can propose an alterative schedule in $\mathcal{T}^*_{conv}(R)$ that makes either groups 1 and 2 or 1 and 3 or 2 and 3 better off. But group 2 has maximum utility under $\hat{t}$ globally. So to prove non-existence we have to construct a schedule that attracts groups 1 and 3.

Before proceeding any further let us note a few points about the parameters of our model.

Parameters of our model

The cut-off $\frac{a_1-a_2}{2a_1-a_2}\mu$ for R tells about whether the resulting $\hat{t}$ is straight or not. (It is straight if R is above it and not straight otherwise.) However, when $\hat{t}$ is not straight, another cut-off for R seems to be important viz $\frac{a_2-a_3}{2a_2-a_3}\mu$. It turns out (as we shall prove shortly) that with $R < \frac{a_2-a_3}{2a_2-a_3}\mu$ (and given that it is already less than $\frac{a_1-a_2}{2a_1-a_2}\mu$), there is no equilibrium while for R in between ($\frac{a_1-a_2}{2a_1-a_2}\mu > R > \frac{a_2-a_3}{2a_2-a_3}\mu$), there could be.

Now the relative position of these cut-offs depends on parameter values and apriori we cannot say which of them is larger than another. $\frac{a_1-a_2}{2a_1-a_2}\mu > \frac{a_2-a_3}{2a_2-a_3}\mu$ if and only if $a_2 < \sqrt{a_1 a_2}$. Similarly for the other inequality. Notice that a_2 less than the geometric mean of the productivity of the other two groups roughly indicates a higher density of population towards lower productivities. The opposite is true for $a_2 > \sqrt{a_1 a_3}$. The first case seems to be closer to reality as far as population in over-populated democracies is concerned. Now the results we enumerate here depends on which of these inequalities is true (for a given population). We summarize this in figure 4.11.

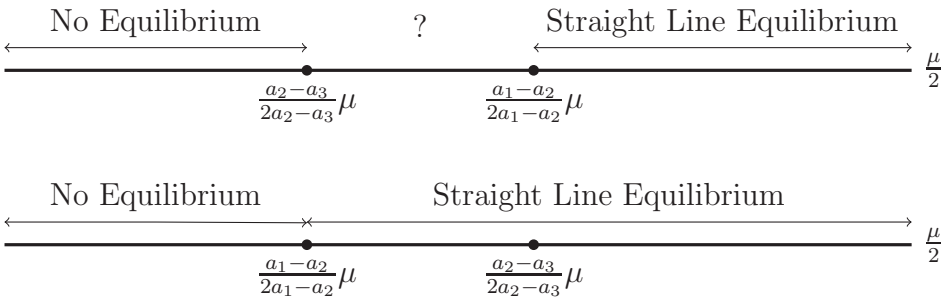


Figure 4.11. Parameters

Theorem 21 *Suppose* $R < \frac{a_1-a_2}{2a_1-a_2}\mu$ *(i.e.* $\hat{t}$ *is not straight) and* $\tan\phi_o < \frac{2(a_2-a_3)}{2a_2-a_3}$. *Then we can find a tax schedule that makes groups 1 and 3 better off than under* $\hat{t}$.

Proof. (For a detailed proof please refer to appendix .7 Taxation-C.) We construct an alternative 3-point tax schedule that makes 1 and 3 pay less taxes and 2 pay more. Suppose we join the origin to the initial P_3 and the take a straight line making angle ρ with the origin with $\rho = \phi_o + d\phi_o$. Then 2 will pay more taxes only if the income it earns under the new schedule in equilibrium is greater than the income that it would earn if it were to pay the same taxes as before under the new schedule. This yields the condition $\tan\phi_o < \frac{2(a_2-a_3)}{2a_2-a_3}$ which holds by assumption. Pick such a ρ.

Under this schedule, revenue collected is $> R$ while 1 and 3 are no worse off (since 1 and 3 pay the same and 2 pays higher). Now shift y_3 downwards along P_3, join this point with the origin thereafter draw it parallel to the previous curve (with angle ρ). Since this curve lies below $\hat{t}$ everywhere, it collects less than R. Hence by continuity of revenue, we can find a curve that collects exactly R and makes 1 and 3 better off. ■

Corollary 22 *When* $R < \frac{a_1-a_2}{2a_1-a_2}\mu$, *then* $R < \frac{a_2-a_3}{2a_2-a_3}\mu$ *is sufficient for no equilibrium to exist.*

(Please see appendix .7 Taxation-C for proof.)

So, to sum up what we've found so far: (i)If a_1, a_2, a_3 are such that $\frac{a_1-a_2}{2a_1-a_2}\mu > \frac{a_2-a_3}{2a_2-a_3}\mu$ (i.e. $a_2 < \sqrt{a_1 a_3}$) then we have shown that equilibrium exists for all $R > \frac{a_1-a_2}{2a_1-a_2}\mu$, it doesn't for all $R < \frac{a_2-a_3}{2a_2-a_3}\mu$.

(ii)If a_1, a_2, a_3 are such that $\frac{a_1-a_2}{2a_1-a_2}\mu < \frac{a_2-a_3}{2a_2-a_3}\mu$ (i.e. $a_2 > \sqrt{a_1 a_3}$) then the second cut-off doesn't matter since it's important only when it's lower than the first one. Hence equilibrium exists for all $R > \frac{a_1-a_2}{2a_1-a_2}\mu$ (and it is a straight line) and doesn't exist otherwise.

Since we have completely characterized existence/non-existence in case the second inequality holds, let us concentrate on what happens if the first inequality holds. The only case to be considered here is then what happens if R lies in between the two cut-offs $\frac{a_1-a_2}{2a_1-a_2}\mu > R > \frac{a_2-a_3}{2a_2-a_3}\mu$. We know that our candidate for equilibrium, $\hat{t}$, is not straight in this range. This brings us to the last step of our road map.

STEP 6

Our hope for $\hat{t}$ to be an equilibrium would be when it is no longer possible to collect R by making 2 worse off and 1 and 3 better off i.e. under the case $\tan\phi_o > \frac{2(a_2-a_3)}{2a_2-a_3}$ (since Theorem 19 tells us that it is possible to make 1 and 3 better off and 2 worse off and collect R if the opposite holds). We do it in steps.

Step 6.A

We first calculate, that given any point $P_2 = (x_2, y_2)$, what is the maximum revenue y_1 that can be collected from group 1. The maximum revenue will be

given when the angle joining P_2 and P_1 satisfies

$$\tan\theta = 1 - \sqrt{\frac{x_2 - y_2}{a_1 - x_2}}$$

Similarly, given any point P_3, the maximum revenue that can be collected from group 2 is given by

$$\tan\phi = 1 - \sqrt{\frac{x_3 - y_3}{a_2 - x_3}}$$

Given this, the maximum tax that can be collected from group 2 can be calculated to be

$$y_2 = (1/2)(\sqrt{a_2 - x_3} - \sqrt{x_3 - y_3})^2 + y_3$$

Step 6.B

Now utility of group 3 at the point $P_3 = (a_3/2, a_3\tan\phi/2)$ can be calculated to be $a_3^2(1 - \tan\phi)/4$ call it $\bar{U}_3$. (For simplification of computation, instead of utility function $\ln(x-y) + \ln(1-x/a)$ we take a monotonic transformation of it, viz $(x-y)(a-x)$.) Hence the indifference curve for group 3 through this point satisfies

$$(x_3 - y_3)(a_3 - x_3) = a_3^2(1 - \tan\phi)/4$$

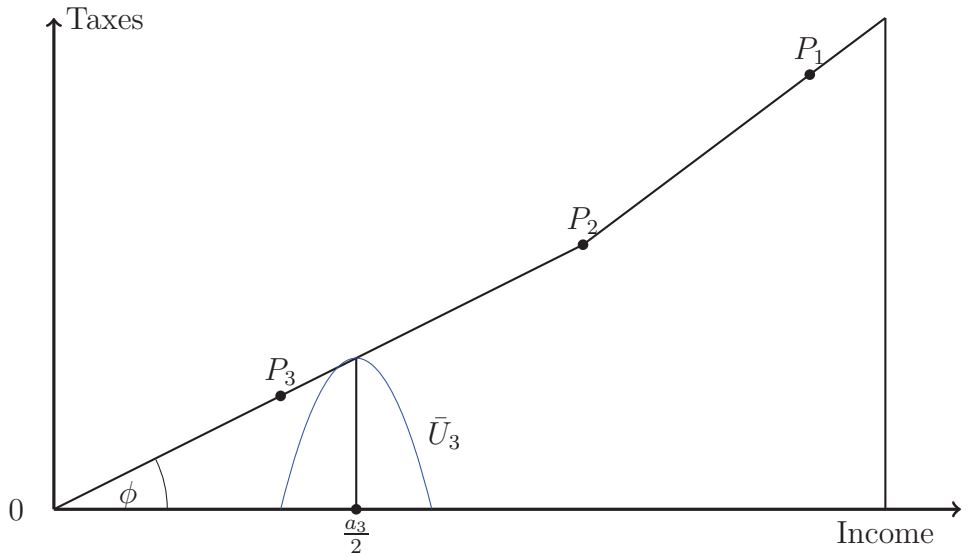


Figure 4.12. Indifference Curve for Group 3

The curve intersects the horizontal axis at two points but only one of them lies below $a_3/2$ which is given by $a_3/2(1-\sqrt{\tan\phi})$. Note that utility is higher on lower indifference curves (since, roughly speaking, taxes to be paid are lower and this can also be formally proved). So to make group 3 better off, we must place him on a lower indifference curve (i.e. $y_3 < (a_3/2)\tan\phi$) and $a_3/2(1-\sqrt{\tan\phi}) \leq x_3 \leq a_3/2$. Substituting y_3 from the indifference curve equation, we can write the maximum collectible revenue from group 2 as a function of x_3 alone.

Step 6.C

Next we want to study how this amount changes as x_3 varies within the relevant range. Differentiating y_2 (expressed in terms of x_3), we get the following

$$\frac{\partial y_2}{\partial x_3} = \frac{1}{2} - \frac{\sqrt{\bar{U}_3(a_2 - x_3)}}{2(a_3 - x_3)^{3/2}} + \frac{\sqrt{\bar{U}_3}}{2\sqrt{(a_2 - x_3)(a_3 - x_3)}} - \frac{\bar{U}_3}{2(a_3 - x_3)^2}$$

$$\implies 2\frac{\partial y_2}{\partial x_3} = 1 - \frac{\bar{U}_3}{(a_3 - x_3)^2} - \frac{\sqrt{\bar{U}_3}(a_2 - a_3)}{(a_3 - x_3)^{3/2}\sqrt{a_2 - x_3}}$$

which shows that as x_3 increases $\frac{\partial y_2}{\partial x_3}$ falls. Now maximum x_3 is $a_3/2$. Also we're considering the case $\tan\phi > \frac{2(a_2-a_3)}{2a_2-a_3}$. Combining the last two facts we get an upper bound on $\bar{U}_3$ given by

$$\bar{U}_3 \leq \frac{a_3^3}{4(2a_2 - a_3)}$$

Using this upper bound and evaluating $\frac{\partial y_2}{\partial x_3}$ at the maximum x_3 we get

$$\frac{\partial y_2}{\partial x_3} \geq 0$$

Therefore, for all other lower x_3's in the range we're interested in, the derivative is strictly positive. Which means y_2 is the maximum possible when x_3 is the maximum possible. Hence any attempt to make 3 better off would result in lower taxes from group 2. So the only possibility left to consider is whether we can propose an alternative that makes 3 better off and collects less taxes from them,

makes 2 worse off and collects less taxes from them, but makes 1 better off and collects more taxes from them, so much more that still compensates the losses from 2 and 3. We prove that in fact, under some more conditions, it is not possible to do so and $\hat{t}$ is the equilibrium in that case.

Step 6.D

The value of utility of group 1 under $\hat{t}$ is given by $(1/(1-\tan\theta))\{(1-\tan\theta)(a_1/2)+(a_2/4)(\tan\theta-\tan\phi)\}^2$, call it $\bar{U}_1$.

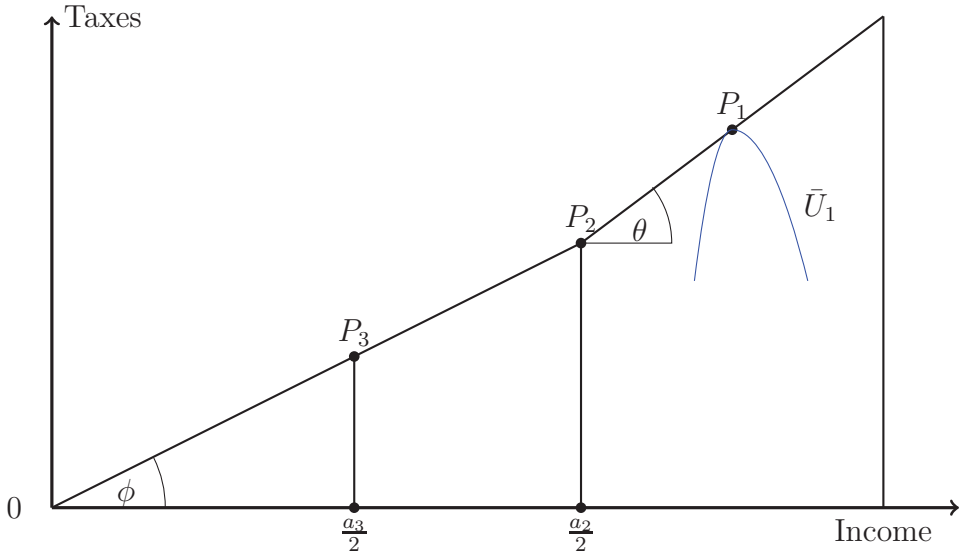


Figure 4.13. Indifference Curve for Group 1

Equation of the indifference curve passing through P_1 (the point where 1 is placed under $\hat{t}$) is given by

$$y = x - \frac{\bar{U}_1}{a_1 - x}$$

Now equation of the tangent to this indifference curve at any point (x_1, y_1) is given by

$$y = x\Big(1 - \frac{\bar{U}_1}{(a_1 - x_1)^2}\Big) + \frac{\bar{U}_1(2x_1 - a_1)}{(a_1 - x_1)^2}$$

If group 3 is given the point $P_3 = (a_3/2, (a_3/2)\tan\phi)$ and a ρ-angled line from that point is given to 2, then 2's utility is maximized at

$$x_2 = \frac{a_2}{2} - \frac{a_3}{2}\frac{\tan\rho - \tan\phi}{2(1-\tan\rho)}$$
$$y_2 = \tan\rho(x_2 - \frac{a_3}{2}) + \frac{a_3}{2}\tan\phi$$

Now if 1's utility were to remain unchanged and if it were to pay higher taxes, then the best possible way to achieve it would be to draw the tangent to 1's $\bar{U}_1$-indifference curve from (x_2, y_2), say at some point (x_1, y_1). (A formal proof is geometrically rather cumbrous and hence is not included here.) Hence (x_2, y_2) would satisfy

$$y_2 = x_2\Big(1 - \frac{\bar{U}_1}{(a_1 - x_1)^2}\Big) + \frac{\bar{U}_1(2x_1 - a_1)}{(a_1 - x_1)^2}$$

Step 6.E

Now revenue from group 3 is unchanged under this allocation compared to $\hat{t}$. So total change in revenue is given by

$$\frac{\partial(\text{Revenue})}{\partial x_1} = \pi_2\frac{\partial y_2}{\partial x_1} + \pi_1\Big(1 - \frac{\bar{U}_1}{a_1 - x_1}\Big)$$

Taking all the above equations into consideration, simple but lenghty calculations yield that under some restriction on the population size, the change in revenue will be negative. That is, for $\tan\phi = \frac{2(a_2-a_3)+\epsilon}{2a_2-a_3}$ (remember we're in the case $\tan\phi > \frac{2(a_2-a_3)}{2a_2-a_3}$), if for some constant, $K(a_1, a_2, a_3, \epsilon)$, that depends on productivities and ϵ (please see appendix .8 Taxation-D for a derivation of K), if $\pi_1 < K\pi_2$ holds, then the loss in revenue cannot be compensated even after making two groups weakly better off.

Hence no alternative tax schedule can be constructed that beats $\hat{t}$, and that will then be our equilibrium. We summarize our finding in the following proposition.

Theorem 23 *Suppose* $\frac{a_1-a_2}{2a_1-a_2}\mu > \frac{a_2-a_3}{2a_2-a_3}\mu$ *(i.e.* $a_2 < \sqrt{a_1 a_3}$*). Suppose* R *is such that* $\frac{a_1-a_2}{2a_1-a_2}\mu > R > \frac{a_2-a_3}{2a_2-a_3}\mu$*. Suppose* $\frac{2(a_1-a_2)}{2a_1-a_2} > \tan\phi_0 > \frac{2(a_2-a_3)}{2a_2-a_3}$*. Then for some appropriate* ϵ*,* $\tan\phi_0$ *can be written as* $\frac{2(a_2-a_3)+\epsilon}{2a_2-a_3}$*. If this* ϵ *also satisfies* $\pi_1 < K\pi_2$*, then* $\hat{t}$ *will be an equilibrium.*

Recall that the upper upper bound for $\tan\phi_0$ comes from $(\tan\phi_0)\mu/2 < R < \frac{a_1-a_2}{2a_1-a_2}\mu$ while the lower bound is the criterion for not being able to collect R by extracting group 2 and making the other two groups better off.

So our kinked equilibrium tax schedule (if there is one) will look like this (figure 4.14):

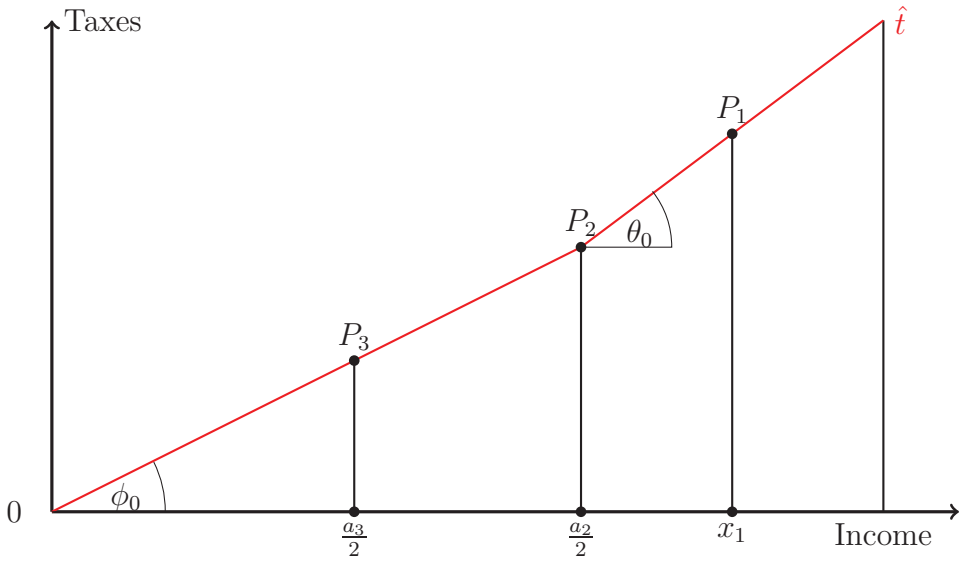


Figure 4.14. Equilibrium Schedule

4.5 The Main Findings : Summary

Recall that there is a one-to-one correspondence between R and ϕ given by the equation

$$\sqrt{2a_1 - a_2} - \sqrt{a_2(1-\tan\phi)} = 2\sqrt{\frac{R - A\tan\phi}{\pi_1}} \quad \ldots\ldots\ldots\ldots\ldots (*)$$

for the tax schedule $\hat{t}$ where A is some function of productivities and population sizes which are the parameters of the model. (Since for every ϕ, we get the tax revenues of groups 2 and 3 and it also gives the maximum collectible revenue from group 1, so that we know what the revenue collected would be. Hence for every R, there's a ϕ that collects it.)

(i) We have shown that for R bigger than some value ($\frac{a_1-a_2}{2a_1-a_2}\mu$), which also means that for $\tan\phi$ bigger than some value $\left(\frac{2(a_1-a_2)}{2a_1-a_2}\right)$, the straight line tax schedule will be the equilibrium out of all convex tax schedules.

(ii) We have shown that for R lower than some number ($\frac{a_2-a_3}{2a_2-a_3}\mu$), which also means that for $\tan\phi$ smaller than some number $\left(\frac{2(a_2-a_3)}{2a_2-a_3}\right)$, there will be no equilibrium out of all convex tax schedules.

Remark 24 *Whether* $\frac{a_1-a_2}{2a_1-a_2}\mu > \frac{a_2-a_3}{2a_2-a_3}\mu$ *or not is not known to us a priori. If it so happens that parameter values are such that* $\frac{a_1-a_2}{2a_1-a_2}\mu < \frac{a_2-a_3}{2a_2-a_3}\mu$*, then the existence/non-existence is completely determined. For all* $R < \frac{a_1-a_2}{2a_1-a_2}\mu$ *there is no equilibrium, and for* R *greater than that the straight line is an equilibrium.*

(iii) For the case $\frac{a_1-a_2}{2a_1-a_2}\mu > \frac{a_2-a_3}{2a_2-a_3}\mu$, suppose our R belongs in between this range (which means our ϕ belongs to $\frac{2(a_1-a_2)}{2a_1-a_2} > \frac{2(a_2-a_3)}{2a_2-a_3}$). Then for some appropriate ϵ, $\tan\phi$ can be written as $\frac{2(a_2-a_3)+\epsilon}{2a_2-a_3}$. If this ϵ also satisfies $\pi_1 < K\pi_2$, then $\hat{t}$ will be an equilibrium. If not, we don't really say anything.

Also notice that the closer a_1 and a_2 are (keeping μ the same), the larger the chances of existence of equilibrium.

In summary, we have shown that by incorporating labor-leisure choice on the part of the voters in an otherwise standard voting-over-taxation model, we could find existence of pure strategy equilibrium in some cases when there was none without incorporating it. We found existence of marginally progressive taxes in the case of

population with larger proportion of less productive people which, by and large, seems to represent large third world democracies.

However there's still a long way to go. For a more complete analysis we must include concave (marginally regressive taxes) in our framework and see whether or not our results are preserved. Our conjecture would be that our existence result will no longer hold in that case. But given this 'social justice'-based support for convex schedules, we think it has been worthwhile looking into such schedules only more closely. And in fact we do find a strictly convex curve as equilibrium schedule (under certain conditions) when in fact the space of schedules considered has been (weakly) convex.

Conjecturing no existence in the more general model, we could perhaps explore minimum additional assumptions that would restore existence back (like parties having 'ideological' stands, citizen-candidate models etc). Also we have assumed majority rule to determine winners. Though this may be representative of political systems/ democracies in a vast part of the globe, yet thinking of other systems of election may also be interesting.

Appendices

.1 Democratization-A

Relation Between Vote Share and the Probability of Winning We establish that vote share and probability of winning move in the same direction.

Proof.

$$\begin{aligned}
&\frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A} = \frac{\partial (P(V_{A,1}(m_A, m_B) \geq \frac{1}{2}))}{\partial m_A} \\
\approx &\frac{P(V_{A,1}(m_A, m_B)|_{m_A + dm_A} \geq \frac{1}{2}) - P(V_{A,1}(m_A, m_B)|_{m_A} \geq \frac{1}{2})}{dm_A} \\
\approx &\frac{P(V_{A,1}(m_A, m_B)|_{m_A} + \frac{\partial V_{A,1}(m_A, m_B)}{\partial m_A}|_{m_A} dm_A \geq \frac{1}{2}) - P(V_{A,1}(m_A, m_B)|_{m_A} \geq \frac{1}{2})}{dm_A}
\end{aligned}$$

It follows that $\frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A}$ will be positive, negative or zero, depending on the sign of $\frac{\partial V_{A,1}(m_A, m_B)}{\partial m_A}$. ■

.2 Democratization-B

Proof of Proposition 1

Proof. Substituting the values of P_2, and W's in (2.3) and (2.4), and differentiating (2.3) w.r.t p_A and (2.4) w.r.t p_B we get the following first order conditions for an interior solution:

$$\begin{aligned}
\frac{\delta + \gamma}{f} =& 4p_A^2 - (p_A + p_B)^2 - 2(1+m)(p_A - p_B) + R(2p_A - (1+m)) \\
\frac{\delta - \gamma}{f} =& -4p_B^2 + (p_A + p_B)^2 - 2(1+m)(p_A - p_B) + R(2p_B - (1+m))
\end{aligned}$$

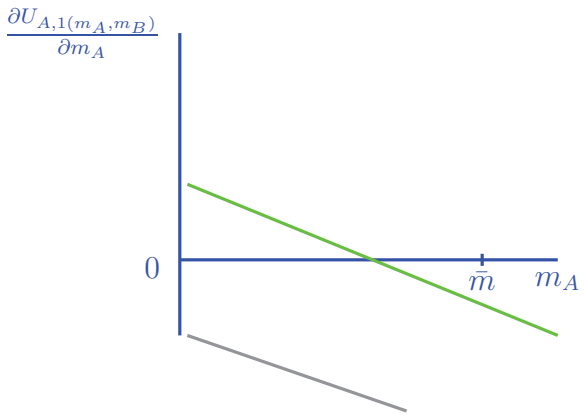
We can solve these to find the above expressions for p_A^* and p_B^*. The conditions in the proposition guarantee that the cumulative distribution functions for v and b are 'interior' (and not 1 or 0 at the relevant values), and that the solutions are indeed interior. ■

.3 Democratization-C

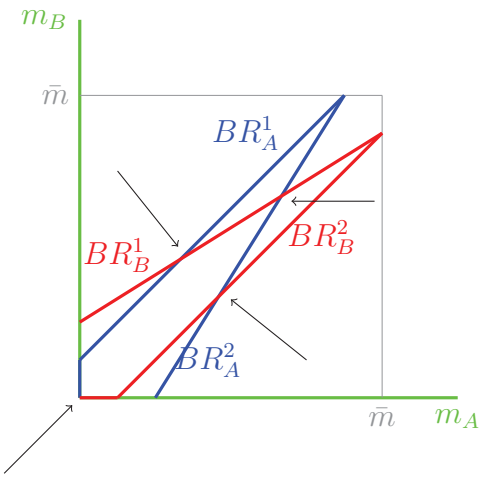
Existence of Nash Equilibrium Proof. Under the conditions of Proposition 2, we can show the following:

$$\frac{\partial^2 U_{A,1}(m_A, m_B)}{\partial m_A^2} < 0$$
$$\frac{\partial U_{A,1}(m_A, m_B)}{\partial m_A}|_{(\bar{m},\bar{m})} < 0$$

This means that, under the assumed conditions, as m_A falls, $\frac{\partial U_{A,1}}{\partial m_A}$ increases, holding m_B constant. Since the latter is negative at $\bar{m}$, it becomes less negative, 0 and then positive as m_A falls (given m_B). Hence we will have something like the following:



Either $\frac{\partial U_{A,1}}{\partial m_A}$ becomes 0 for some $m_A \in (0, \bar{m})$ or it becomes 0 for $m_A < 0$, so that the best response of A is 0. Also we can show that $\frac{\partial^2 U_{A,1}(m_A,m_B)}{\partial m_A \partial m_B}$ is > 0. This means that as m_B falls, $\frac{\partial U_{A,1}}{\partial m_A}$ also falls, so that optimal m_A also falls (or stays at 0). Now we know that at $m_B = \bar{m}$, A's optimal m_A is below $\bar{m}$ (as described above). Hence we have A's best responses that might look like BR_A^1 or BR_A^2 (or identically 0). Similarly, we will have B's possible best response start at $m_B < \bar{m}$ for $m_A = \bar{m}$. Hence the best response of A will be a vertical line, that of B a horizontal line, and they are bound to intersect, either at an interior or at 0. Thus we will have the following:



In any case, a Nash equilibrium exists, given the nature of the best responses under the conditions assumed. ■

.4 Democratization-D

Brief Proof of Proposition 2

Proof. We have the following:

$$\frac{\partial U_{A,1}(m_A, m_B)}{\partial m_A} = \text{Direct Effect for A} + \text{Indirect Effect for A}$$
$$\frac{\partial U_{B,1}(m_A, m_B)}{\partial m_B} = \text{Direct Effect for B} + \text{Indirect Effect for B}$$

where

$$\text{Direct Effect for A: } \frac{\partial P_{A,1}(m_A, m_B)}{\partial m_A}(R + U_{A,2}(m_A) - U_{A,2}(m_B))$$
$$\text{Direct Effect for B: } \frac{\partial P_{B,1}(m_A, m_B)}{\partial m_B}(R + U_{B,2}(m_B) - U_{B,2}(m_A))$$

And the indirect effects can be computed by substituting the solutions from Proposition 1 back in (2.3) and (2.4) and differentiating with respect to m_A and

m_B respectively, to get the following:

$$\text{Indirect Effect for A: } P_{A,1}(m_A, m_B)\frac{\partial U_{A,2}(m_A)}{\partial m_A} = \frac{1}{2}P_{A,1}(m_A, m_B)$$
$$\text{Indirect Effect for B: } P_{B,1}(m_A, m_B)\frac{\partial U_{B,2}(m_B)}{\partial m_B} = -\frac{1}{2}P_{B,1}(m_A, m_B)$$

Now $\frac{\partial P_{A,1}(m_A,m_B)}{\partial m_A}$ and $\frac{\partial P_{B,1}(m_A,m_B)}{\partial m_B}$ are negative as long as $\int_{\bar{m}}^{1} \frac{\partial L_j}{\partial m}$ is positive. Substituting the values of P_2, u_j's etc and simplifying, we get

$$\frac{\partial U_{A,1}(m_A, m_B)}{\partial m_A} \sim -\frac{\delta}{4\gamma} + \frac{Rf}{4\gamma}(\bar{m} - m_A) + \frac{1}{4}.$$

The expression on the right hand side is negative if

$$(\bar{m} - m_A) < \frac{\delta - \gamma}{Rf}.$$

Moreover, $\frac{\partial U_{B,1}(m_A,m_B)}{\partial m_B}$ is < 0 whenever $\frac{\partial U_{A,1}(m_A,m_B)}{\partial m_A}$ is < 0 (notice the negative indirect effect for B so that extension is more natural for B). Hence if A extends up to m_A (given m_A satisfies the above condition), then B will extend up to below m_A. Hence this Proposition states extension up to 'at least m', which is basically the condition for A extending up to m (i.e. $m_A = m$ in the above condition), since B is extending even further.

Putting $m_A = 0$ in the above the above condition, we see that A extends up to 0 if

$$\bar{m} < \frac{\delta - \gamma}{Rf}.$$

Moreover, from interiority conditions of the policies (asymptotically), we need

$$\bar{m} < \frac{Rf - \delta - \gamma}{Rf}.$$

Hence we get

$$\bar{m} < \min\{\frac{\delta - \gamma}{Rf}, \frac{Rf - \delta - \gamma}{Rf}\}.$$

■

.5 Women's suffrage-A

Results by dropping missing data

As noted in chapter 3, about 30% of the data on competition (% of Democratic seats in Upper and Lower houses and % of Republican seats in Upper and Lower houses) are absent. They are either missing or are absent because there was no legislative session in that year for that state. Either way, it might be helpful to get a sense of what the results would be if only the available observations were included in the data i.e. for the results in this appendix, all the absent observations were dropped (not interpolated or kept at the last available value). We reestimate the three basic models with this new data.

Table A-1: Basic model

Model 1[a]	Coefficient	Robust Standard Error	p-value	Marginal Effect
Competition dummy	1.3490	1.1325	.234	.0120
Sex-ratio (female/male)***	−.0683	.0162	.000	−.0003
% of urban population	.0329	.0231	.154	.0001
% of black population	−.0838	.0520	.108	−.0004
% of female working	−.0775	.0676	.251	−.0003
Constant	1.8153	1.3692	.185	
Log likelihood	−60.2183			
Number of observations	1052			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. The marginal effect of a continuous variable, x, on the dependent variable, y, is computed as the slope $\frac{\partial y}{\partial x}$, at the sample mean of x and holding all the other variables constant at either their sample means (or other specific values). The marginal effects of the dummy variables are based on switches from zero to one, holding all else constant at sample means (or other specific values). The probability of suffrage for model (1) was computed at the sample means of all the variables and it is about .004. Marginal effects are also calculated at this probability.

Table A-2: Model with 'diffusion'

Model 2[a]	Coefficient	Robust Standard Error	p-value	Marginal Effect
Competition dummy	1.7048	1.0937	.119	.0188
Sex-ratio (female/male)***	−.0628	.0191	.001	−.0003
% of urban population***	.0692	.0190	.000	.0003
% of black population	−.0606	.0523	.247	−.0003
% of female working*	−.1659	.0735	.024	−.0007
'Diffusion' dummy***	2.3814	.5038	.000	.0158
Constant	−.1535	1.9848	.938	
Log likelihood	−53.2788			
Number of observations	1052			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability of about .004 (at which the marginal effects would be calculated) was computed for model (2) by letting the 'diffusion dummy' equal .42 (not equal to its sample mean of .147) and holding others at their sample means.

Table A-3: Model with 'diffusion' and 'duration dependence'

Model 3[a]	Coefficient	Robust Standard Error	p-value	Marginal Effect
Competition dummy*	2.3168	1.3215	.080	.0370
Sex-ratio (female/male)***	−.1285	.0323	.000	−.0006
% of urban population***	.1197	.0327	.000	.0005
% of black population*	−.1977	.1152	.086	−.0009
% of female working***	−.4782	.1607	.003	−.0021
'Diffusion' dummy**	1.444	.5980	.016	.0113
Duration dependence***	.1518	.0510	.003	.0007
Constant*	5.1845	3.0020	.084	
Log likelihood	−46.8384			
Number of observations	1052			

[a]*, **, and *** indicate significance at 90%, 95% and 99% respectively. To facilitate comparison across models, the same probability, like in model (1) of about .004 (at which the marginal effects would be calculated) was computed for model (3) by letting 'duration dependence' equal 42 (not equal to its sample mean of 19.5946) and holding others at their sample means.

Compared to the results where missing observations were not dropped, we can make the following observations:

1. The number of observations is 1052 (from 1665), so the additional ones were interpolated.

2. The number of non-zero outcomes (of full suffrage dummy) is now 13 (from 16), so three of them were corresponding to missing values of the competition data.

3. The p-value of the competition dummy variable has increased and so has the size of its coefficient.

4. The log-likelihood has greatly increased in all the models (they were around -80 to -70 previously).

5. However the number of clusters has reduced to 47 (from 48). That's because of Utah. It is there in the data set for just one year- 1896. It enters statehood in 1896 and gives women's suffrage the same year. But the competition variable values are missing for that year. So now (with the new data set) Utah is no longer there. Previously (with interpolated competition variable values) I had competition values interpolated from previous years and had Utah in the data set for one year.

Comparing the elasticities of the continuous variables across the models, we get the following:

Table A-4: Comparing elasticities of continuous variables

Elasticities[a]	Model 1	Model 2	Model 3
Sex-ratio (female/male) in %	−6.3512	−5.8447	−11.9599
% of urban population	1.1736	2.4680	4.2716
% of black population	−.8635	−.6248	−2.0387
% of female working	−1.2152	−2.6014	−7.4993

[a]To facilitate comparison across models, the elasticities were computed for the same probability of about .004 in all the models by choosing the 'diffusion' and 'duration dependence' variables in models (2) and (3) suitably, as described in the footnotes of tables A-2 and A-3.

.6 Taxation-A

Convex Tax schedules

Let t be a convex, twice differentiable tax schedule. All results in this section concern such a tax schedule. Let also $U = \ln(C)+\ln(1-L)$ where $C = aL-t(aL)$ denotes consumption. Fix productivity a.

Claim 25 *There exists an $L \in [0,1]$ which maximizes U in the sense that it solves the FOC.*

Proof. The FOC for maximization is given by

$$\frac{a(1-t'(aL))}{aL-t(aL)} = \frac{1}{1-L}$$

Given t, we want to know whether we can find an L for every a that solves the FOC. If there does exist such an $L(a)$, let $aL(a) = x(a)$.Then the FOC becomes

$$\frac{a(1-t'(x(a)))}{x(a)-t(x(a))} = \frac{1}{1-L(a)}$$

Hence if there is a x solving the above equation, so is there an L (since $L = x/a$). The above condition is

$$\frac{1-t'(x(a))}{x(a)-t(x(a))} = \frac{1}{a-x(a)}$$
$$\implies a = x + \frac{x(a)-t(x(a))}{1-t'(x(a))}$$

Call the RHS function $f(x)$. Since t is twice differentiable, t' and t are continuous and hence f as a function of x is continuous. We see that $f(0) = 0$ and $f(a) \geq a$ (since $x-t(x) \geq 0$ and $t'(x) < 1$ for all x). Hence there exists $x \in [0,a]$ such that $f(x) = a$ and therefore an L that solves the FOC. ■

Claim 26 *Higher the productivity (a), higher is the earned income ($aL(a)$) in equilibrium.*

Proof. Differentiating $f(x)$ from the proof of the last claim, we get

$$f'(x) = 2 + \frac{t''(x)(x-t(x))}{(1-t'(x))^2}$$

which is > 0 since $t'' > 0$, t being convex, and $t(x) < x$, by assumption on tax schedules. ($f'(x)$ is not just positive, it is strictly greater than 2.) Hence $f(x)$ is strictly increasing in x and since $f(x) = a$, $a_1 > a_2 \iff x_1 > x_2$ where $x_i = a_i L(a_i)$. ■

Claim 27 *The maximum labor supplied, for any a is $\leq 1/2$ in equilibrium.*

Proof. Now $x = aL(a)$. Hence $(dx/da) = L + a(dL/da)$. Also $a = f(x)$ so that $1 = f'(x)(dx/da)$. Noting that $f'(x)$ is > 2, $(dx/da) < (1/2)$ which implies $a(dL/da) < (1/2 - L)$. Since a is non-negative, we get that in equilibrium, if $L \geq 1/2$ for some a, then for all higher a, optimal labor is lower.

Now suppose that the maximum labor supplied for any given a correspond to a_0 with $L(a_0) > 1/2$. Take $a_1 = a_0 - \delta$ with δ small so that $L(a_1) > 1/2$ (this can be done due to the continuity of L as a function of a). Then $a(dL(a_1)/da)|_{a_1} < (1/2) - L(a_1)$ which is < 0. But then, for a higher a, say a_0, optimal labor must be lower i.e. $L(a_0) < L(a_1)$ a contradiction to $L(a_0)$ being maximum labor supplied. Therefore, maximum labor supplied must be $\leq 1/2$. ■

Proposition 28 *Let $\mathbf{U} = \ln(\mathbf{C}) + \ln(\mathbf{1} - \mathbf{L})$. Let $\sum_{i=1}^{3} \pi_i a_i = \mu$. Then the maximum revenue that can be collected through a convex, twice differentiable tax schedule in $\mathcal{T}_{\mathbf{conv}}$ is $\mu/\mathbf{2}$.*

Proof. Let income earned by group i be x_i. Since maximum possible labor, L_i, in equilibrium is $1/2$, maximum possible earned income is $a_i L_i = x_i = a_i/2$. Hence

$$\begin{aligned}
\text{Revenue} &= \sum_{i=1}^{3} \pi_i t(x_i) \\
&\leq \sum_{i=1}^{3} \pi_i x_i \\
&\leq \sum_{i=1}^{3} \pi_i \frac{a_i}{2} \\
&= \mu/2
\end{aligned}$$

■

Notice that μ is like the mean of the productivities. So this proposition tells us that with convex tax schedules we can collect atmost half the average income as taxes.

Remark 29 *Though the value '$\mu/2$' maybe specific to the utility function chosen, we think similar claims will hold for all utility functions characteristically similiar to the one we've chosen.*

Remark 30 *If the revenue to be collected is $> \mu/2$, then no convex tax schedule would serve the purpose. Hence if $R > \mu/2$, the only feasible tax schedules to consider are concave, the marginally regressive ones. This is in sharp contrast with the exogenous income case where any $R < \mu$ could be collected through an appropriate convex tax schedule.*

Remark 31 *Though we have considered only twice differentiable convex functions, we believe we can prove the same result for any convex tax schedule.*

For illustration, we construct a concave schedule that collects more than $\mu/2$.

A Concave Tax Schedule

We still retain the assumption $U = \ln(C) + \ln(1 - L)$. But now let t be the following:

$$t(x) = \begin{cases} (\tan\phi)x & \text{if } x < \frac{a_2}{2} \\ \frac{a_2}{2}\tan\phi + \tan\theta(x - \frac{a_2}{2}) & \text{if } x \geq \frac{a_2}{2} \end{cases}$$

with $0 < \theta < \phi < 1$. Notice that the resulting tax schedule is concave.

For proof that this tax schedule, for an appropriate choice of the angles θ and ϕ, collects more than $\mu/2$, see Appendix Taxation-B.

Under this tax schedule the taxes that each of the groups pay are the following:

$$\begin{aligned} y_1 =& \tan\theta\Big(\frac{a_1}{2} - \frac{a_2(\tan\theta - \tan\phi)}{4(1-\tan\theta)} - \frac{a_2}{2}\Big) + \frac{a_2}{2}\tan\phi \\ y_2 =& \frac{a_2}{2}\tan\phi \\ y_3 =& \frac{a_3}{2}\tan\phi \end{aligned}$$

Now we know that for every ϕ, the θ that gives maximum revenue is given by the equation (look at Step 5 of Section 4)

$$1 - \tan\theta = \sqrt{\frac{a_2(1-\tan\phi)}{2a_1 - a_2}}$$

Hence after substituting for y_1, y_2, y_3 and $\tan\theta$, the total possible revenue under this schedule is given by

$$(\pi_2 a_2 + \pi_3 a_3)\frac{\tan\phi}{2} + \frac{\pi_1 a_1}{2} + \frac{\pi_1 a_2 \tan\phi}{4} - \frac{\pi_1}{2}\sqrt{a_2(2a_1 - a_2)(1-\tan\phi)}$$

As $\phi \to 1$, revenue $\to \mu/2 + \pi_1 a_2/4$ which is $> \mu/2$.

.7 Taxation-B

Proof of Lemma 12

Proof. Equation of the P_2P_3 segment is given by

$$y - y_3 = \frac{y_2 - y_3}{x_2 - x_3}(x - x_3) = \tan\rho(x - x_3)$$

Given our utility function, the FOC for maximization yields

$$\frac{(1-\tan\rho)}{x - y_3 - x\tan\rho + x_3\tan\rho} = \frac{1}{a_2 - x}$$

Manipulating we get

$$x = \frac{a_2}{2} - \frac{x_3\tan\rho - y_3}{2(1-\tan\rho)}$$

Since $y_3 = x_3\tan\phi$, we have

$$x = \frac{a_2}{2} - x_3\frac{\tan\rho - \tan\phi}{2(1-\tan\rho)}$$

For group 2 to choose P_2 when given the P_2P_3 segment, $x_2 \leq x$ which gives the right hand inequality of the second condition in the lemma. Similarly we can get the other inequalities.

Suppose group 2 is given the P_1P_2 segment, then its utility will be maximized at

$$x = \frac{a_2}{2} - \frac{x_1 \tan\theta - y_1}{2(1 - \tan\theta)}$$

Therefore, for group 2 to choose P_2 when given the P_1P_2 segment, $x_2 \geq x$ which gives the left hand inequality of the second condition in the lemma. ■

.8 Taxation-C

Proof of Theorem 19

Proof. We want to find a condition such that 2 gives more taxes when ρ increases slightly from ϕ_0, given 1 and 3 are kept at the same position. Under angle ρ group 2 pays taxes according to $y = \tan\rho(x - a_3/2) + a_3 \tan\phi_0/2$. This yeilds utility maximizing

$$\tilde{x}_2 = \frac{a_2}{2} - \frac{a_3}{4}\frac{\tan\rho - \tan\phi_0}{1 - \tan\rho}$$

Computations yield that if revenue is unchanged under the old and the new schedules then

$$x_2 = \bar{x}_2 = \frac{a_3}{2} + \frac{(a_2 - a_3)\tan\phi_0}{2\tan\rho}$$

So revenue is higher if and only if

$$\tilde{x}_2 > \bar{x}_2$$

Taking $\rho = \phi_0 + d\phi$ and taking first order approximations we see that this holds if

$$\tan\phi_0 < \frac{2(a_2 - a_3)}{2a_2 - a_3}$$

Since $\tan\phi_0 < 2R/\mu$, we see that

$$R < \frac{(a_2 - a_3)}{2a_2 - a_3}\mu$$

is sufficient for the above condition to hold. ■

.9 Taxation-D

Calculation of $K(a_1, a_2, a_3, \epsilon)$

Let

$$x_2 = \frac{a_2}{2} - \frac{a_3}{2}\frac{\tan\rho - \tan\phi}{2(1-\tan\rho)} \tag{8}$$

$$y_2 = \tan\rho(x_2 - \frac{a_3}{2}) + \frac{a_3}{2}\tan\phi \tag{9}$$

$$y_2 = x_2\Big(1 - \frac{\bar{U}_1}{(a_1 - x_1)^2}\Big) + \frac{\bar{U}_1(2x_1 - a_1)}{(a_1 - x_1)^2} \tag{10}$$

$$\tag{11}$$

We first calculate $\frac{\partial y_2}{\partial x_1}$ from equations (9), (10), (11). Differentiating equation (10) we get

$$\frac{\partial y_2}{\partial x_1} = \tan\rho\Big(\frac{\partial x_2}{\partial x_1}\Big) + \frac{\partial\tan\rho}{\partial x_1}\Big(x_2 - \frac{a_3}{2}\Big) \tag{12}$$

Differentiating (9) we get

$$\frac{\partial x_2}{\partial x_1} = -\frac{a_3(1-\tan\phi)}{4(1-\tan\rho)^2}\frac{\partial\tan\rho}{\partial x_1} \tag{13}$$

$$\implies \frac{\partial\tan\rho}{\partial x_1} = -\frac{\partial x_2}{\partial x_1}\frac{4(1-\tan\rho)^2}{a_3(1-\tan\phi)} \tag{14}$$

Substituting (14) in (12) we get

$$\frac{\partial y_2}{\partial x_1} = \frac{\partial x_2}{\partial x_1}\Big(\tan\rho - \Big(x_2 - \frac{a_3}{2}\Big)\frac{4(1-\tan\rho)^2}{a_3(1-\tan\phi)}\Big) \tag{15}$$

Now given $\tan\rho > \tan\phi$ and we're in the case with $\tan\phi > \frac{2(a_2-a_3)}{2a_2-a_3}$, we can verify that the expression in the parenthesis of (15) is non-negative. Differentiating (11) we get

$$\frac{\partial y_2}{\partial x_1} = \frac{\partial x_2}{\partial x_1}\Big(1 - \frac{\bar{U}_1}{(a_1-x_1)^2}\Big) + \frac{2\bar{U}_1}{(a_1-x_1)^3}(x_1-x_2) \tag{16}$$

Substituting $\frac{\partial x_2}{\partial x_1}$ from (16) into (15) we get

$$-\frac{\partial y_2}{\partial x_1} = \frac{2\bar{U}_1(x_1-x_2)}{(a_1-x_1)^3}\frac{S}{1-\frac{\bar{U}_1}{(a_1-x_1)^2}-S} \tag{17}$$

where $S = \tan\rho - \Big(x_2 - \frac{a_3}{2}\Big)\frac{4(1-\tan\rho)^2}{a_3(1-\tan\phi)}$. Now the denominator of the RHS of (17) is also positive since slope of the tangent decreases as x_1 increases from $\bar{x}_1$. That is, $\frac{\bar{U}_1}{(a_1-x_1)^2} \leq \frac{\bar{U}_1}{(a_1-\bar{x}_1)^2} = 1-\tan\theta$ and we can verify $(\tan\theta - S) > 0$ which together imply $\Big(1 - \frac{\bar{U}_1}{(a_1-x_1)^2} - S\Big) > 0$. Hence $-\frac{\partial y_2}{\partial x_1} > 0$.

Since $\frac{\partial \text{Revenue}}{\partial x_1} = \pi_2\frac{\partial y_2}{\partial x_1} + \pi_1\Big(1 - \frac{\bar{U}_1}{(a_1-x_1)^2}\Big)$, for the change in revenue to be negative, we need

$$\frac{\pi_1}{\pi_2} < \frac{\Big|\frac{\partial y_2}{\partial x_1}\Big|}{1-\frac{\bar{U}_1}{(a_1-x_1)^2}} \tag{18}$$

However it is enough to have $\frac{\pi_1}{\pi_2} < \Big|\frac{\partial y_2}{\partial x_1}\Big|$. Hence we now try to find a lower bound for $-\frac{\partial y_2}{\partial x_1}$ so that $\frac{\pi_1}{\pi_2}$ lower than that bound will be our sufficient condition.

$$\begin{aligned}
-\frac{\partial y_2}{\partial x_1} &= \frac{2\bar{U}_1(x_1-x_2)}{(a_1-x_1)^3}\frac{S}{1-\frac{\bar{U}_1}{(a_1-x_1)^2}-S} \\
&\geq \frac{2\bar{U}_1(x_1-x_2)}{(a_1-x_1)^3}\frac{S}{1-\frac{\bar{U}_1}{(a_1-x_1)^2}} \\
&\geq \frac{2\bar{U}_1(x_1-x_2)}{(a_1-x_1)^3}S \\
&\geq \frac{2\bar{U}_1(\bar{x}_1-\frac{a_2}{2})}{(a_1-\bar{x}_1)^3}\Big(\tan\phi\Big(\frac{2a_2}{a_3}-1\Big) - 2\Big(\frac{a_2}{a_3}-1\Big)\Big)
\end{aligned}$$

Now let $\tan\phi = \frac{2(a_2-a_3)+\epsilon}{2a_2-a_3}$. Then the expression in the parenthesis becomes $\frac{\epsilon}{3}$. $\bar{U}_1$ and $\bar{x}_1$ have $\tan\theta$ in their expressions and $\tan\theta$ can be found by plugging in $\tan\phi$ in $(1-\tan\theta) = \sqrt{\frac{a_2(1-\tan\phi)}{2a_1-a_2}}$. Hence the lower bound for $-\frac{\partial y_2}{\partial x_1}$ reduces to some number that depends on a_1, a_2, a_3 and ϵ. Call it K. Since all the terms

determining K are positive we get $-\frac{\partial y_2}{\partial x_1} \geq K > 0$. So, let for the desired change in revenue, $\pi_1 < K\pi_2$ hold.

References

[1] Acemoglu, Daron and Robinson, James A.; A Theory of Political Transitions; The American Economic Review, (2001) Vol. 91, No. 4; Pages 938- 963

[2] Acemoglu, Daron and Robinson, James A.; Economic Origins of Dictatorship and Democracy; Cambridge University Press (2005)

[3] Acemoglu, Daron and Robinson, James A.; Why did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective; The Quarterly Journal of Economics, (2000) Vol. 115, No 4; Pages 1167-1199

[4] Aidt, Toke and Jensen, Peter S.; Tax Structure, Size of Government, and the Extension of the Voting Franchise in Western Europe, 1860-1938; Forthcoming in International Tax and Public Finance

[5] Aidt, Toke and Jensen, Peter S.; The Taxman Tools Up: An Event History Study of the Introduction of the Personal Income Tax, 1815-1938; Forthcoming in the Journal of Public Economics

[6] Allison, Paul D.; Event History Analysis: Regression for Longitudinal Event Data; Sage Publications (1984)

[7] Alston, Lee J., Dupre, Ruth and Nonnenmacher, Tomas; Social Reformers and regulation: the prohibition of cigarettes in the United States and Canada; Explorations in Economic History (2002); Vol. 39, Pages 425-445

[8] Atkinson, A.B., Public Economics in Action: The Basic Income/ Flat Tax Proposal (1995), Clarendon Press, Oxford

[9] Atkinson, Neill; Adventures in Democracy: A History of the Vote in New Zealand; University of Otago Press (2003)

[10] Bertocchi, Graziella; The Enfranchisement of Women and the Welfare State; C. E. P. R Discussion Paper (2007)

[11] Blossfeld, Hans-Peter; Golsch, Katrin; Rohwer, Gotz; Event History Analysis with Stata; Lawrence Erlbaum Associates, Publishers (2007)

[12] Blossfeld, Hans-Peter; Hamerle, Alfred; Mayer, Karl Ulrich; Event History Analysis: Statistical Theory and Applications in Social Sciences; Lawrence Erlbaum Associates, Publishers (1989)

[13] Boehmke, Frederick J. and Witmer, Richaed; Disentangling Diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion; Political Research Quarterly (2004), Vol. 57, No. 1, Pages 39 - 51

[14] Box-Steffensmeier, Janet M.; Jones, Bradford S.; Event History Modeling: A Guide for Social Scientists; Cambridge University Press (2004)

[15] Bradford, Jones S., and Branton, Regina P.; Beyond Logit and Probit: Some Alternative Duration Modeling Strategies for State Policy Adoption, Preprint 2003

[16] Buckley, Jack and Westerland, Chad; Duration Dependence, Functional Form, and Corrected Standard Errors: Improving EHA Models of State Policy Diffusion; State Politics and Policy Quarterly (2004), Vol. 4, No. 1, Pages 94 - 113

[17] Carbonell-Nicolau, O and Ok, E.A., Voting over Income Taxation, Journal of Economic Theory, forthcoming

[18] Carmines, Edward G., The Mediating Influence of State Legislatures on the Linkage Between Interparty Competition and Welfare Policies, The American Political Science Review (1974), Vol. 68, Pages 1118-1124

[19] Case, Anne C., Rosen, Harvey S., and Hines, James R., Jr; Budget spillovers and fiscal policy interdependence: Evidence from the states; Journal of Public Economics (1993), Vol. 52, Pages 285-307

[20] Caul, Miki; Political Parties and the Adoption of Candidate Gender Quotas: A Cross National Analysis; The Journal of Politics (2001), Vol. 63, No. 4, Pages 1214-1229

[21] Chen, Anthony S and Phinney, Robin; Did the Civil Rights Movement Have a Direct Impact on Public Policy? Evidence from the Passage of State Fair Housing Laws, 1959-1965; Working Paper No. 04-005, Gerald R. Ford School of Public Policy (2004)

[22] Chen, Anthony S.; The Passage of State Fair Employment Legislation, 1945-1964: An Event History Analysis with Time-Varying and Time-Constant Covariates; Institute of Industrial Relations Working Paper Series, Working Paper #79 (2001)

[23] Collier, Ruth Berins; Paths Toward Democracy: The Working Class and Elites in Western Europe and South America; Cambridge University Press (1999)

[24] Conley, John P. and Temimi, Akram; Endogenous Enfranchisement When Groups' Preferences Conflict; Journal of Political Economy, (2001) Vol. 109, No. 1; Pages 79- 102

[25] Crawford, Finla Goff; State Government; Henry Holt and Company, Inc., New York (1932)

[26] Cukierman, A, Meltzer, A.H., and Richard, S.F., Political Economy, (1991), Oxford University Press

[27] Dahan, M and Strawczynski, M, Optimal Income Taxation: An example with a U-Shaped Pattern of Optimal Marginal Tax Rates: Comment (2000), The American Economic Review (Vol 90 No.3)

[28] Dahl, Robert A. ed; Political Oppositions in Western Democracies; Yale University Press (1966)

[29] Daley, Caroline and Nolan, Melanie ed; Suffrage and Beyond: International Feminist Perspectives; New York University Press, Washington Square, New York (1994)

[30] Diamond, Larry; Linz, Juan J. and Lipset, Seymour Martin ed; Politics in Developing Country: Comparing Experiences with Democracy, Lynne Reinner Publishers (1990)

[31] Diamond, P.A., Optimal Income Taxation: An example with a U-Shaped Pattern of Optimal Marginal Tax Rates (1998), The American Economic Review (Vol 88 No.1)

[32] Doepke, Matthias and Tertilt, Michele; Women's Liberation: What's in It for Men? Preprint 2007

[33] Doyle, William R.; Adoption of Merit-Based Student Grant Programs: An Event History Analysis; Educational Evaluation and Policy Analysis (2006), Vol. 28, No. 3, Pages 259-285

[34] Dye, Thomas R.; Politics, Economics, and the Public: Policy Outcomes in the American States; American Politics Research Series (1966)

[35] Fishback, Price V. and Kantor, Shawn Everett.; The Adoption of Workers' Compensation in the United States 1900-1930; Journal of Law and Economics, Vol. XLI (1998); Pages 305-341

[36] Flexner, Eleanor; Century of Struggle: The Woman's Rights Movement in the United States; The Belknap Press of Harvard University Press (1959)

[37] Ghosal, Sayantan and Proto, Eugenio; Enfranchisement, Intra-Elite Conflict and Bargaining; The Warwick Economics Research Paper Series (TWERPS) (2006)

[38] Goldin, Claudia; Understanding the Gender Gap: An Economic History of American Women; Oxford University Press (1990)

[39] Golembiewski, Robert T.; A Taxonomic Approach to State Political Party Strength; Western Political Quarterly (1958); Pages 494-513

[40] Gray, Virginia; Hanson, Russell L.; Jacob, Herbert; Politics in the American States: American States: A Comparative Analysis; C Q Press, Washington, D.C (1999)

[41] Grimshaw, Patricia; Women's Suffrage in New Zealand; Auckland University Press, Oxford University Press (1972)

[42] Grossman, Gene M. and Helpman, Elhanan; Special Interest Politics; The MIT Press (2001)

[43] Hettich, W and Winer, S. L., Democratic Choice and Taxation (1999), Cambridge University Press

[44] Himmelfarb, Gertrude; The Politics of Democracy: The English Reform Act of 1867; Journal of British Studies, (1966) VI, Pages 97 138

[45] Hofferbert, Richard I. And Sharkansky, Ira; State and Urban Politics; Readings in Comparative Public Policy; Little, Brown and Company, Boston (1971)

[46] Hofferbert, Richard I., The Relation between Public Policy and some Structural and Environmental Variables in the American States, The American Political Science Review (1966); Pages 73-82

[47] Jack, William and Lagunoff, Roger; Dynamic Enfranchisement; Journal of Public Economics, Vol. 90 (2006), Pages 551-572

[48] Jewell, Malcolm E.; The State Legislature: Politics and Practice; Random House Studies in Political Science, New York (1962)

[49] Keyssar, Alexander; The Right to Vote: The Contested History of Democracy in the United States; Basic Books, A Member of the Persus Books Group (2000)

[50] Khan, Zoria B; Married Women's Property Laws and Female Commercial Activity: Evidence from United States Patent Records, 1790 -1895; The Journal of Economic History (1996), Vol. 56, No. 2, Pages 356 - 388

[51] King, Brayden G.; Cornwall, Marie and Dahlin, Eric C; Winning Woman Suffrage One Step at a Time: Social Movements and the Logic of the Legislative Process; Social Forces (2005), Vol. 83, No. 3, Pages 1211-1234

[52] Klor, E. F., On the Popular Support for Progressive Taxation (2003), Journal of Public Economic Theory, 5 (4), pp. 593-604

[53] Landes, William M. and Solomon, Lewis C.; Compulsory Schooling Legislation: An Economic Analysis of Law and Social Change in the Nineteenth Century; The Journal of Economic History (1972), Vol. 32, No. 1, Pages 54-91

[54] Lipset, Seymour Martin; The Political Man: The Social Bases of Politics; Doubleday & Company, Inc., Garden City, New York (1960)

[55] Lizzeri, Alessandro and Persico, Nicola; Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an application to Britain's "Age of Reform"; Quarterly Journal of Economics, (2004) Vol. 119(2), Pages 707-765

[56] Llavador, Humberto and Oxoby, Robert J.; Partisan competition, growth, and the franchise,' Quarterly Journal of Economics, (2005) Vol. 120, Pages 1155-1188

[57] Marhuenda, F, Ortuno-Ortin, I, Popular Support for progressive taxation (1995), Economics Letters 48, pp. 319-324

[58] Matthaei, Julie A.; An Economic History of Women in America: Women's Work, the Sexual Division of Labor, and the Development of Capitalism; Schocken Books (1982)

[59] McCammon, Holly J. and Campbell, Karen E.; Winning the Vote in the West: The Political Successes of the Women's Suffrage Movements, 1866-1919; Gender and Society; Volume 15 No. 1, February 2001; Pages 55-82

[60] McDonagh, Eileen L. and Price, H. Douglas; Woman Suffrage in the Progressive Era: Patterns of Opposition and Support in Referenda Voting, 1910-1918, The American Political Science Review, (1984) Vol. 79, Pages 415 435

[61] Miller, Grant; Women's Suffrage, Political Responsiveness, and Child Survival in American History; preprint February 2008

[62] Mirrlees, J.A., An exploration in the Theory of Optimum Income Taxation (1971), Review of Economic Studies, Vol. 38, No. 2, pp. 175-208

[63] Mitra, T and Ok, E. A., Personal Income Taxation and the Principle of Equal Sacrifice Revisited (1996), International Economic Review, Vol. 37, No. 4, pp 925-948

[64] Mitra, T, Ok, E. A. and Kockesen, L, Popular support for progressive taxation and the relative income hypothesis (1998), Economics Letters 58, pp. 69-76

[65] Moore, Barrington (Jr); Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World; Beacon Press, Boston (1966)

[66] Munger, Frank (edited); American State Politics: Readings for Comparative Analysis; Thomas Y. Crowell Company, New York (1966)

[67] Oldfield, Audrey; Woman Suffrage in Australia: A Gift or a Struggle?; Cambridge University Press (1992)

[68] Pande, Rohini; Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India; The American Economic Review (2003); Pages 1132-1151

[69] Persson, Torsten and Guido, Tabellini; Political Economics: Explaining Economic Policy; The MIT Press (2000)

[70] Ranney, Austin and Kendall, Willmoore; The American Party Systems; American Political Science Review (1954); Pages 477-485

[71] Rincke, Johannes; Policy Diffusion in space and time: The case of charter schools in California school districts; Regional Science and Urban Economics (2007), Vol. 37, Pages 526-541

[72] Roemer, J.E., The Democratic Political Economy of Progressive Income Taxation (1999), Econometrica, Vol. 67, No. 1, pp. 1-19

[73] Rotella, Elyce J.; From Home to Office: U.S. Women at Work, 1870-1930; UMI Research Press (1981)

[74] Rustow, Dankwart A.; The Politics of Compromise: A Study of Parties and Cabinet Government in Sweden, Princeton University Press (1955)

[75] Salanie, B., The Economics of Taxation, (2002), The MIT Press

[76] Salomone, William A.; Italy in the Giolittian Era: Italian Democracy in the Making, 1900-1914, University of Pennsylvania Press (1945)

[77] Schlesinger, Joseph A.; A Two-Dimensional Scheme for Classifying the States According to Degree of Inter-party Competition; American Political Science Review (1955); Pages 1120-1128

[78] Scott, Anne Firor and Scott, Andrew Mackay; One Half the People: The Fight for Woman Suffrage; University of Illinois Press (1982)

[79] Shefter, Martin; Regional Receptivity to Reform: The Legacy of the Progressive Era; Political Science Quaterly (1983); Vol. 98, No. 3, Pages 459 - 483

[80] Skocpol, Theda; The Enactment of Mothers' Pensions: Civic Mobilization and Agenda Setting or Benefits of the Ballot?: Response; The American Political Science Review (1995)

[81] Smith, Daniel A., and Fridkin, Dustin; Delegating Direct Democracy: Interparty Legislative Competition and the Adoption of the Initiative in the American States; American Political Science Review; Vol. 102 No. 3; Pages 333-350 (2008)

[82] Tentler, Leslie Woodcock; Wage-Earning Women: Industrial Work and Family Life in the United States, 1900-1930; Oxford University Press (1979)

[83] The Complete Report by the Citizens Conference on State Legislatures; State Legislatures: An Evaluation of Their Effectiveness; Praeger Publishers (1971)

[84] Ticchi, Davide and Vindigni, Andrea; On Wars and Political Development. The Role of International Conflicts in the Democratization of the West; preprint October 2005

[85] Tuomala, M., On optimal non-linear income taxation: numerical results revisited, preprint

[86] Tuomala, M., Optimal Income Tax and Redistribution (1990), Clarendon Press, Oxford

[87] Valenzuela, Arturo; Political Brokers in Chile: Local Government in a Centralized Polity, Duke University Press (1977)

[88] Verney, Douglas V.; Parliamentary Reform in Sweden 1866-1921, Oxford Clarendon Press (1957)

[89] Walker, Jack L.; The Diffusion of Innovations among the American states; The American Political Science Review (1969); Vol. 63, Pages 880 -899

[90] Webster, R. A.; Industrial Imperialism in Italy 1908-1915, University of California Press (1975)

[91] Yamaguchi, Kazuo; Event History Analysis; Applied Social Research Methods Series, Volume 28; Sage Publications (1991)

Curriculum Vita

Soumyanetra Munshi

2004 - 2009 Ph. D. in Economics, Rutgers University

2004 - 2006 Master of Arts in Economics, Rutgers University

2002 - 2004 Master of Science in Quantitative Economics, Indian Statistical Institute, Calcutta, India

1999 - 2002 Bachelor of Science in Economics, Presidency College, University of Calcutta, Calcutta, India

2008 - 2009 Graduate School Teaching Assistantship in the Writing Program of the department of English, Rutgers University

2005 - 2008 Graduate School Teaching Assistantship in the department of Economics, Rutgers University