Public versus Private Provisioning: Role of Education and Political Participation*

Mausumi Das
Delhi School of Economics

Adway De
Delhi School of Economics

Skand Goel
Delhi School of Economics

Tridip Ray†
Indian Statistical Institute, Delhi

This draft: 4 December 2013

Abstract

This paper studies the role played by education in the public provision of private 'merit goods', such as healthcare, schooling, security and so on. Corruption is endemic in public provision. Better educated individuals are more effective at exerting political pressure, which reduces corruption and improves quality of the merit goods delivered. At the same time, educated elite have higher income which allow them to opt out of public provisioning and form a private club that delivers the merit good/service to its members. This may lead to deterioration of public provisioning. Depending on parametric conditions, several equilibrium configurations exist, some exhibiting multiple equilibria – with different degrees of corruption and concomitant variation in the quality of public provision and welfare of people. Under a stochastic adaptive dynamic process, almost surely a unique equilibrium will be selected, which need not be the one which is *least corrupt* or *most efficient*. This brings in the scope for effective policy intervention. We also analyze the long run wealth dynamics and its implication for the public vis-a-vis private provisioning.

JEL Classification: H42, I24, I25, P16, P26.

Keywords: Education, political participation, public provision, private provision.

*For helpful discussions and suggestions we thank Abhijit Banerji, Parikshit Ghosh and Anirban Kar. All remaining errors are ours.

†Corresponding author. Contact Address: Economics and Planning Unit, Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, Qutab Institutional Area, New Delhi 110016, India. Email: `tridip@isid.ac.in`

1 Introduction

Public delivery of essential services is one of the major functions of the welfare state and forms the backbone of daily sustenance of masses. Especially in low income countries, government provision of healthcare, schooling, infrastructure and public security is instrumental in reducing vulnerability of poorer populations. Despite the buoyant economic growth in the emerging economies in recent decades, there has not been significant improvement in the quality of these services. Their delivery is often marred by wastefulness and corruption (see, for example, Delavallade (2006)). Thus 'public action', civic participation and building political pressure are crucial in bringing the state to function better.

There are divergent views about how civic participation has shaped up in recent years, especially in the developing countries. For example, consider the case of India. The recent mass demonstrations against corruption, often concentrated in urban areas, has been interpreted by many to represent a resurgence of the public conscience as people came out on the streets to demand more accountability and efficiency from the state. However, at the same time, several observers worry that certain sections of the society have remained apathetic towards the political process. These factors have obviously impacted upon the effectiveness of the state provisioning. While the state of public healthcare is clearly less than desirable, the surge in "corporate" hospitals in the recent years presents a stark contrast. Inequality, especially in consumption, appears to have only been accentuated. While there have been some improvements in provision of merit goods, there is also the perception that those who can afford to are resorting to greater dependence on market instead of government provisioning. The sudden burgeoning of integrated townships near metropolitan cities in the last few years is a striking example of what appears to be a more general trend. If the "motto of public service is inclusion, expansion and efficiency" (Bhattacharya et al., 2012), then such a trend may in fact be deprecatory as it reduces the incentives for a section of society to engage in the political process.

How does education come into this scenario? The role of education is not limited to mere skill formation and the resulting increase in wage earnings. Education impacts upon one's overall personality – by molding one's preferences, by inculcating new values, by making one more conscious of her rights and more articulate in voicing her opinion. All these factors have implications for the choices that a person makes in the spheres of economics, politics and other social activities. In this paper we focus on one such aspect that links education to political activism. Observational studies have found strong correlation between education and political participation, voting, and civic awareness at the individual level (see Nie (1996) for a survey). Putnam (1995) claimed that educational attainment is "best individual level predictor of political participation". While the evidence from recent studies has been mixed,[1] it is still not unreasonable to argue that educational attainment leads to facility in the public sphere in general and political efficacy in particular. This may be more emphatically true in developing countries where informational asymmetry and socioeconomic inequalities have greater bite.[2]

This paper builds a simple model to study the effect of education and educational inequality on the public provision of private merit goods and services, such as healthcare, schooling, security and infrastructure. We start with the presumption that educated individuals are more effective in creating political pressure to ensure better delivery of public services. As political action is a public good, it creates a positive externality for the uneducated. However, if the anticipated service provision is not satisfactory, then the educated individuals may also opt out of the public process and form a private club that delivers the merit good only to its members. This would lead to increased "corruption" and deterioration of government provisioning. We analyze the conditions under which the educated elite opts out of public provisioning in order to form a club and its impact on the quality of public services. In our model, the agents are differentiated by wealth levels which leads to differential educational attainment. In this setup, we find that the stage game throws up several possibilities depending on the parametric conditions, many of which exhibit multiple equilibria with varying degrees of corruption and concomitant quality of public provisioning. To study the equilibrium selection, we posit a stochastic adaptive process such that the economy (almost surely) gravitates towards a unique equilibria. However, which of the multiple equilibria will be selected depends on the initial conditions as well as on the parametric configurations. In particular, a 'club' equilibrium (characterized by the educated elite opting out of public provision, resulting in high corruption and low quality of public provisioning) is more likely when the initial educational inequality is large. Moreover poor quality of education, which

[1] For instance, Milligan et al. (2004) find sizable effects of schooling on voting behavior in the U.S., but much weaker effects in the U.K. and Germany. Chevalier and Doyle (2012) survey this literature and present fresh evidence to argue that US is in fact an outlier in the relationship between schooling and voting behavior. Stepan et al. (2011) document a negative relationship between voting propensity and educational attainment in India. Milligan et al. (2004) and find that additional years of high school significantly increase interest in politics, efforts to acquire information about political issues/campaigns, and beliefs in freedom of speech.

[2] In the context of colonial Benin, Wantchekon et al. (2013) find that students were significantly more likely to campaign for political parties, or even become full-fledged members. Their findings show a clear effect of education on political participation. They claim that these are the first (quasi) experimental evidence in support of the positive effect of education on political participation in developing countries.

reduces the effectiveness of education in the political arena, and/or a repressive political regime, which increases the cost of participating in political activism, increase the likelihood of the club equilibrium. On the other hand, legal and institutional barriers that increase the fixed cost of club formation would increase the chances of a 'public' equilibrium (characterized by the educated elite staying under public provision, resulting in low corruption and high quality of public provisioning). Since some of these parameters are amenable to policy interventions, the model allows us to qualitatively comment on policy mechanisms that result in realization of the low corruption public equilibrium.

Surprisingly, low corruption need not necessarily be welfare enhancing. In fact, when we rank these two equilibria in terms of welfare of various groups of agents, we find that under certain scenarios, the high corruption club equilibrium would Pareto-dominate the low corruption public equilibrium. This apparently counter-intuitive result is explained by the fact that when some educated elite choose to leave public provisioning to join the club, aggregate political effort to curb corruption indeed goes down, but congestion in public provisioning is reduced as well. Thus, depending on the number of educated elite who leave public provisioning, it is conceivable that the de-congestion effect dominates the corruption effect so that per capita provision for those who remain in public provisioning goes up improving their welfare.

We then extend the static choice of merit good provision to a dynamic set up where the distribution of wealth and the consequent educational inequality evolve endogenously from the historically given initial conditions. We find that the adverse effects of educational inequality noted in the static model get amplified in the dynamic context: if the initial inequality is high enough, then in the long run the economy will inevitably reach a steady state where the entire educated elite joins the club leaving the quality of public provisioning at its lowest.

The rest of the paper is structured as follows. Section 2 presents some supporting evidence for the assumption that education increases effectiveness in political activity and discusses the related literature. Section 3 sets up the basic model. Individual choices are analyzed in section 4. Section 5 explains the possibilities of multiple equilibria and equilibrium selection. In section 6 we derive some comparative static results and discuss their economic and political implications. Section 7 discusses welfare ranking of equilibria and highlights the possibility that a low corruption equilibrium may not always be the one with higher welfare. Section 8 analyzes the long run dynamics. Finally, section 9 concludes the paper.

2 Motivational Evidence and Related Literature

The primary assumption of our model is that educated individuals are more effective at political action. In this section we motivate this assumption by referring to various supporting evidence from the literature. The idea that education reduces both the cognitive and material costs of political participation is quite dominant in the political science literature (Wolfinger and Rosenstone, 1980). Informed decision making is cognitively costly and education reduces that cost. Moreover, low levels of education can magnify the costs of access to information for decision making (this is especially true for developing countries). Brady et al. (1995) find that among other things, education affects individuals' participation through improved civic skills and political knowledge. Using a slightly broader array of outcomes as indicators of democratic citizenship, Nie (1996) finds similar results. Using an instrumental variables approach, Dee (2004) finds that additional schooling appears to increase the quality of civic knowledge as measured by the frequency of newspaper readership. In a field experiment in Kenya, Friedman et al. (2011) find that young women's participation in a scholarship program that has boosted enrollment rates in Kenya made it more likely for them to read newspapers and, as a result, acquire more objective knowledge about politics and express less satisfaction with Kenya's democracy and current economic conditions. Popkin and Dimock (1999) show that citizens with low levels of information cannot follow public discussion of issues, are less accepting of the give and take of democratic policy debates, make judgments on the basis of character rather than issues, and are significantly less inclined to participate in politics at all. More educated individuals can also better monitor corruption and are more likely to blow the whistle (Botero et al., 2012). Indeed, in surveying the literature, Galston (2001) notes that there are indications of an emerging consensus on the notion that competent democratic citizens need not be policy experts, but there is a level of basic knowledge below which the ability to make a full range of reasoned civic judgments is impaired.

Another explanation that has gained currency with the recent surge of interest in social capital is that educated people have better social networks and display higher trust. This facilitates collective action leading to greater political efficacy (Nie, 1996; Helliwell and Putnam, 2007). A similar notion underlies the model of Glaeser et al. (2007). Third, in certain societies, education may also affect participation because politicians and organizations tend to target educated citizens in their efforts to mobilize participation. Reciprocally, the educated, with arguably better organizational capabilities and resources, may be in a better position to mobilize the masses (rather than the other way round). Hirschman (1978) noted that the threat of exit may itself bolster political voice. There is evidence that the threat of exit is more credible if it comes from a more educated populace (Warren, 2011). Better political decision making often requires adopting a long term view. Education increases patience (Perez-Arce, 2011), potentially leading to better political choices. Education is also seen as politically empowering. Uslaner and Rothstein (2012) show a powerful statistical link between education levels in 1870 and corruption levels in 2010 for 78 countries and go on to argue that more equal societies educated more of their citizens, which then gave their citizens more opportunities and power, reducing corruption. Basu and King (2001) find that educated Bangladeshi women are more likely to participate in political meetings and to speak up.

Lastly, it is crucial to appreciate that education, as sociologists have long argued, is also a socialization mechanism (Bourdieu, 1986). Political voice is likely to be more effective if both the solicitor and the solicited "speak the same tongue", enabling what Taylor (1997) calls "politics of recognition". In other words, education helps level the playing field to some extent. Appadurai (2004) seems to share this sentiment when he writes that "for voice to take effect, it must engage social, political, and economic issues in terms of ideologies, doctrines, and norms which are widely shared and credible, even by the rich and powerful." Education, therefore, can enhance political voice in multiple ways.

More generally, our paper relates to two broad streams of literature - one focusing on the impact of education on developmental outcomes through the political economy channels; the other dealing with public provision of private goods. The first set of literature explores the link between education and political institution. The idea that education encourages and strengthens democracy has a long intellectual history going back at least to Dewey (1916) and Lipset (1959), the latter tracing the basic idea to Aristotle. The argument of what has come to be known as the "modernization hypothesis" is that education weakens subscription to traditional attachments in favor of merit, fosters preference for democratic values and creates a populace conducive to democracy.[3] More recently, a number of studies have explored the role of education in the choice of political institution (democracy vis-a-vis

[3]The validity of causality from education to democracy implied by the modernization view is not an empirically settled issue. Cross country studies, starting with Barro (1999) who finds primary schooling to be an important predictor of democracy, have yielded contrasting evidence (Glaeser et al., 2004; Acemoglu et al., 2005; Castelló-Climent, 2008) . Papaioannou and Siourounis (2008) look at 174 countries to study democratization in the period 1960-2005 and find that that education significantly increases the probability, intensity, and speed of democratization. Using historical time series, Murtin and Wacziarg (2011) show that that primary schooling, more so than GDP per capita, and more so than secondary and tertiary education, has been a major factor in the democratic transition over 1870-2000. For a revised version of the modernization hypothesis and corroborating evidence see Inglehart and Welzel (2009).

oligarchy) and its implication for development and growth (see, for example, Bourguignon and Verdier (2000), Eicher et al. (2009), Botero et al. (2012), Fortunato and Panizza (2011), Glaeser et al. (2007) etc.) The work from this literature that comes closest to our paper is Campante and Chor (2012a), who justify their empirical finding that political participation is more responsive to schooling in land-abundant countries, and less responsive in human capital-abundant countries in terms of a theoretical model. They propose a time allocation (human capital) model[4] where both the level of political participation and its responsiveness to schooling increase with the relative abundance of the factor that is used in the least skill intensive sector, in this case land.[5] In their model, the government succeeds in extracting a proportion of the citizens' productive income. This proportion, similar to our model, is a decreasing function of the aggregate political effort of citizens. However, their paper does not allow the agents to opt out of the political arena altogether and build an alternative institution as a substitute. This is precisely what we do in our paper.

The second strand of literature that our paper relates to deals with public provision of private goods, starting with Stiglitz (1974).[6] His results pertain to voting over tax rates where tax revenues go towards increasing the quality of public provision. Stiglitz showed that when a private alternative exists, the preference over tax rate is not single peaked. Intuitively, non-single-peakedness occurs because, at low levels of public service quality, a household that prefers high-quality service may prefer the private alternative. Moderate increases in quality from a low base may make the household worse off because taxes rise while the increase in service quality is not sufficient to induce the household to consume the public alternative. Large increases in public service quality, by contrast, may make the household better off. Epple and Romano (1996) and Glomm and Ravikumar (1998) develop this result further and, among other things, prove results relating to existence of majority voting equilibrium. Lülfesmann and Myers (2011) build on Epple and Romano (1996) and evaluate what they label "the slippery slope argument" in the context of voting. The argument is as follows: "giving rich elites the opportunity to opt out will trigger a political process leading to lower taxes and a lower quality of public services; this partial collapse, will make the public system less desirable with less political support which then could lead to a further deterioration of

[4]Ehrlich and Lui (1999) also propose a time allocation model. In their model, however, bureaucrats use their accumulated political capital for collecting more corruption rents.

[5]They propose a similar explanation for the recent democratic uprisings in the Middle East in Campante and Chor (2012b).

[6]There is another literature that tries to understand normative reasons such as equity (Besley and Coate, 1991) and efficiency (Fang and Norman, 2008) for public provisioning of private goods that is not of immediate relevance to us.

the public system." This is very similar to the situation we are studying, albeit not in the case of voting but with respect to direct political effort. In all these models, however, the private good is bought from the market. We on the other hand focus on 'club' provisioning where the agents who opt out form a private club that caters to their (and only their) needs.

There are two papers in this literature (Helsley and Strange (1998) and Bhattacharya et al. (2012)) which consider about public vis-a-vis club provisioning as competing institutions. Helsley and Strange (1998) study the competition between the public sector and a "private government", a voluntary organization that provides only its members with a supplement to public provision at a cost to the members only. The private government cares about aggregate welfare of its members, while the public sector maximizes the aggregate welfare of the entire population. They find that the presence of a private government induces a reduction in public service provision, a result which is similar to ours. Bhattacharya et al. (2012) construct a game theoretic model where the government supplies a public good but where the rich, if dissatisfied, can get together and form a club to supplement that provision. They find that a welfare maximizing government will reduce tax rate in case of club formation, much in the spirit of the 'slippery slope argument'. While these latter two papers are close in spirit to our work, we explicitly link the choice between two competing form of institutions and associated trade offs to the level education of an agent and existing educational inequality. In this sense, our paper links these two broad sets of literature within a unified framework.

3 The Model

At any point of time, the economy has N agents who live for one period. In the next period, an exact replica is born to each agent, who carries the dynastic link forward.[7] Agents differ in their initial wealth they receive from their parents, with the wealthy individuals holding a *gross* wealth x^e and poor ones holding x^u.

We assume education to be a binary variable, i.e., an agent is either educated or uneducated. Whether to educate a child is decided by the parent. Depending on this education decision, the *net* wealth inherited by an agent i is $x_i - e$ if educated, and x_i if uneducated, where e is the cost of education. The educated agents (E) earn a wage y_e and the uneducated agents (U) earn y_u, with $y_e > y_u$. There is a proportional tax, τ, on wage. Therefore, the

[7]In what follows, we study the static game and bring in the temporal aspect only in section 8. That is where we explicitly introduce time subscripts.

income of the agent depends on his education level and is given by

$$\tilde{y}_i = \begin{cases} (x_i - e) + y_e(1-\tau) & \text{if educated,} \\ x_i + y_u(1-\tau) & \text{if not educated.} \end{cases} \tag{1}$$

All agents obtain utility from the consumption of a normal good, c, "warm glow" bequests, b, and from the consumption of another private "merit good" or service, G, such as healthcare, schooling, security, public infrastructure, and so on. G is the focus of our paper. This commodity could be provided by the government (*public provision*) or by individuals forming a private club (*club provision*). Before moving on to the preference structure, we describe each type of provision of G.

First, we consider *public provisioning* of G, G^{pub}. Let $Y \equiv n^e y_e + (N - n^e) y_u$, where n^e is the number of educated people in the economy, be the aggregate income of the economy. Therefore, total tax revenues are τY. This tax revenue, τY, is the potential resource for public provision of G, but the corrupt government may not utilize this resource to its full potential unless forced by political efforts of the agents. The collective political effort is the sum of the individual political efforts, $P_i \in [0, \bar{P}]$, of each individual i: $\sum_{i \in E} \lambda P_i + \sum_{i \in U} P_i$.[8] We assume that the effectiveness of an educated individual's private effort is higher. $\lambda \in (1, \bar{\lambda}]$ is a measure of effectiveness of education in political action. This effect of education, which has been discussed at length in the introduction, is a crucial assumption of our model. We assume the following functional form for publicly provided G,

$$G^{pub} = \frac{\sum_{i \in E} \lambda P_i + \sum_{i \in U} P_i}{N \bar{P} \bar{\lambda}} \cdot \tau Y.$$

G^{pub} is the total amount of public provision. The constants $N\bar{P}\bar{\lambda}$ in the denominator are used for normalization of 'collective political effort', such that $0 \leq \frac{\sum_{i \in E} \lambda P_i + \sum_{i \in U} P_i}{N \bar{P} \bar{\lambda}} \leq 1$. The interpretation of the above formulation is that a proportion $(1 - \frac{\sum_{i \in E} \lambda P_i + \sum_{i \in U} P_i}{N \bar{P} \bar{\lambda}})$ of tax revenues are lost to corruption as long as some individuals are uneducated or some individuals do not participate to the fullest in the political process. The formulation implies that a full recovery of tax revenue is possible only if all individuals are educated, lend maximum political effort, $\bar{P}$, and the nature of education is such that it is extremely effective in making people's political voice strong, i.e., $\lambda = \bar{\lambda}$.

[8] Note that in the present formulation there is no complementarity between political efforts of the individuals. We discuss the implications of this assumption later.

The 'merit goods' we consider are rival in nature. If I have occupied a hospital bed or a seat in a primary school, then you may not. Thus every individual gets only a fraction of the total G^{pub},

$$g^{pub} = \frac{G^{pub}}{n^{pub}} = \frac{1}{n^{pub}} \left(\frac{\sum_{i \in E} \lambda P_i + \sum_{i \in U} P_i}{N \bar{P} \bar{\lambda}} \cdot \tau Y \right), \tag{2}$$

where n^{pub} is the number of people availing of public provision. In other words, collective political action determines the per capita provision of the service.

Political effort is costly. We assume the costs to be quadratic, $C(P_i) = \frac{\delta P_i^2}{2}$. δ can be given several interpretations. One interpretation is that a higher δ indicates a more repressive regime - a more repressive regime imposes higher costs on political action.

Now consider *club provision* of G, G^{club}. Let $n^{club} \equiv N - n^{pub}$ be the number of people opting for club provision. A club is formed when individuals get together to contribute to the cost of G^{club}. This cost has two components: a fixed component F,[9] which is shared by the n^{club} members making the effective per capita burden F/n^{club}, and a variable component for each individual. Each individual i decides how much to contribute towards club provision, g_i. Moreover, we assume F to be sunk or waste. This gives

$$G^{club} = \sum_{i=1}^{n^{club}} g_i - F.$$

Due to the rival nature of the good, as in the case of public provisioning, the individual consumption of G^{club} is given by

$$g^{club} = \frac{G^{club}}{n^{club}} = \frac{\sum_{i=1}^{n^{club}} g_i - F}{n^{club}}. \tag{3}$$

Now we describe the individual choice setting. The agents have identical preferences that are quasilinear in g, the per capita provision of G,

$$U_i(c, b, G) = \alpha \log c_i + (1 - \alpha) \log b_i + g - C(P_i) \tag{4}$$

with $0 < \alpha < 1$.

There are three choices each agent makes. First, an agent chooses whether he wants to go for public provision or club provision. Second, contingent on the first choice, he chooses his optimal actions (consumption, bequest and political effort or club contribution). Moreover,

[9]The interpretation of this fixed cost is similar to the one in Bhattacharya et al. (2012) and Helsley and Strange (1998). Moreover, our club contribution scheme is closer to Helsley and Strange (1998) as compared to Bhattacharya et al. (2012) as the latter model the club as imposing a proportional tax on its members.

an agent must decide whether he wants to educate his child (if he can afford to do so). We discuss each of these next.

The first choice pertains to G. An individual may choose to opt for club provision or public provisioning of G, both of which have been described above. One cannot opt for a mix of the two. If the fixed cost, F, for setting up the club is too high and the income of the poor, $\tilde{y}_u$, too low, the poor agents never find it optimal to join the club. We impose the following condition that ensures that the poor never join the club:[10]

$$\log\left(1-\frac{\frac{F}{n^e}}{\tilde{y}_u}\right)-\frac{F}{n^e}<\frac{1}{\delta}\left(\frac{\tau Y}{N\bar{P}\bar{\lambda}n^u}\right)^2\left(n^u-\frac{1}{2}\right)-\tilde{y}_u+n^e. \tag{5}$$

Contingent on the choice of G, the agent decides to allocate his income between consumption, bequest and club contribution (g_i) as well as chooses his optimal level of political effort. Moreover, the decision to educate his child rests with the agent. He may allocate a part of his warm glow bequest towards his child's education if the bequest is greater than e, the cost of education.

4 Individual Choices

In this section, we proceed by solving for individual choices backwards. First, we look at the education decision of the parent. Then we look at optimal consumption, bequest, political efforts and/or club contributions of individuals (educated and uneducated) conditional on the type of provisioning they opt for. Finally, in the following section, we compare the indirect utilities arising from these choices in order to characterize the best response of an individual in choosing between public and club provision.

4.1 Education Decision

Given his bequest, educating his child is like a portfolio decision for the parent. In choosing whether to do so, a parent considers only the income of the child that would result from education. An individual will choose to educate his child only if $(b-e)+y_e(1-\tau)>b+y_u(1-\tau)$, where b is the optimal bequest. This requires

$$e<(y_e-y_u)(1-\tau). \tag{6}$$

[10]This is similar to Bhattacharya et al. (2012). See Appendix A.3 for derivation of the condition.

We assume (6) holds. Therefore, every parent who can afford education prefers to educate his child. That is, agent i chooses to educate if $b_i = x_i \geq e$. We further assume that

$$x^u < e < x^e. \tag{7}$$

With this assumption, the number of educated individuals is simply n^e and the number of uneducated individuals is $n^u = N - n^e$.

4.2 Public Provisioning

Now, we look at the second stage of choices. First, consider an individual who has opted for public provisioning. The optimal club contribution of such a person is $g_i^* = 0$. The income is given by (1). Therefore, the problem faced by such an agent is

$$\begin{aligned} &\max_{c_i, b_i, P_i} \alpha \log c_i + (1-\alpha)\log b_i + \tfrac{1}{n^{pub}}\left(\frac{\sum_{i\in E}\lambda P_i + \sum_{i\in U} P_i}{N\bar{P}\bar{\lambda}} \cdot \tau Y\right) - \tfrac{\delta P_i^2}{2} \\ &\text{subject to } \; c_i + b_i \leq \tilde{y}_i, \end{aligned} \tag{8}$$

where $\tilde{y}_i$ is given by (1). Solving for the optimal choices (see Appendix A.1 for details) we get that the optimal choice of political effort for an educated individual who has decided to opt for public provisioning is

$$P_i^{e*} = \frac{\lambda \tau Y}{\delta N \bar{P}\bar{\lambda} n^{pub}}, \tag{9}$$

and for an uneducated individual in public provisioning is

$$P_i^{u*} = \frac{\tau Y}{\delta N \bar{P}\bar{\lambda} n^{pub}}. \tag{10}$$

The optimal consumption and bequests are given by

$$\begin{aligned} c_i^* &= \alpha \tilde{y}_i\,, \\ b_i^* &= (1-\alpha)\tilde{y}_i. \end{aligned}$$

Note that both P_i^{e*} and P_i^{u*} are decreasing in n^{pub}. As the number of people opting for public provision increases (decreases), the gains from an individual's effort accrue to a larger (smaller) number of people while the entire cost is still borne by the individual, leading to a free-rider effect.

Given the condition 5 stated above, the uneducated always opt for public provisioning. This gives us $n^{pub} = n^u + (n^e - n^{club})$. Plugging (9) and (10) in (2), we get

$$g^{pub} = \frac{\left(\lambda^2\left(n^e - n^{club}\right) + n^u\right)}{\delta} \cdot \left(\frac{\tau Y}{(n^u + n^e - n^{club})N\bar{P}\bar{\lambda}}\right)^2. \tag{11}$$

An increase in the n^{pub} has two opposite effects on G^{pub}. The "provisioning effect" increases the number of people soliciting public provision, leading to an increase in provision. At the same time, owing to the "free rider effect", the effort of each individual also falls. The latter effect, however, is a second order effect. Therefore, the net provisioning effect on the total amount of public provision, G^{pub}, is positive. There is an additional "congestion effect" on g^{pub} as the amount of public provision available *per capita* also falls. The positive net provisioning effect effect dominates initially, while the congestion effect eventually takes over. This makes g^{pub} rise initially with an increase in n^{pub}, reach the peak at some $\tilde{n}^{club}$ and then fall.[11]

Plugging the optimal values of c_i, b_i and P_i, and $\tilde{y}_i$ (from (1)) into (4), we get the indirect utility function for an educated individual opting for public provision:

$$\begin{aligned} V_e^{pub}(x_i, n^{club}; \lambda, \delta, \tau, n^u) = \ & A + \log\left[(x_i - e) + y_e(1-\tau)\right] - \tfrac{\delta}{2}\left(\frac{\lambda\tau Y}{\delta N\bar{P}\bar{\lambda}(n^u + n^e - n^{club})}\right)^2 \\ & + \left(\frac{(\lambda^2\left(n^e - n^{club}\right) + n^u)}{\delta}\right)\left(\frac{\tau Y}{(n^u + n^e - n^{club})N\bar{P}\bar{\lambda}}\right)^2, \end{aligned}$$

where $A = \alpha\log\alpha + (1-\alpha)\log(1-\alpha)$ is a constant. Similarly, the indirect utility of an uneducated individual is given by

$$\begin{aligned} V_u^{pub}(x_i, n^{club}; \lambda, \delta, \tau, n^u) = \ & A + \log\left[x_i + y_u(1-\tau)\right] - \tfrac{\delta}{2}\left(\frac{\tau Y}{\delta N\bar{P}\bar{\lambda}(n^u + n^e - n^{club})}\right)^2 \\ & + \left(\frac{(\lambda^2\left(n^e - n^{club}\right) + n^u)}{\delta}\right)\left(\frac{\tau Y}{(n^u + n^e - n^{club})N\bar{P}\bar{\lambda}}\right)^2. \end{aligned}$$

We now look at the behavior of V_e^{pub} as a function of n^{club}. Differentiating V_e^{pub} with respect to n^{club} it can be checked that $V_e^{pub}(x_i, n^{club}; .)$ reaches a maximum at $\hat{n}^{club} = n^e - 1 - n^u\left(1 - \frac{2}{\lambda^2}\right)$. A look at the expressions for g^{pub} and V_e^{pub} tells us that besides the opposing "net provisioning effect" and "congestion effect" on g^{pub} as discussed above, there is another effect that operates on the indirect utility with a change in n^{pub}: an increase in n^{pub} (or an equivalent fall in n^{club}, all else fixed) reduces the political cost incurred by an individual because the optimal political effort falls. The combined effect gives V_e^{pub} the inverted-U shape: rising initially with an increase in n^{club}, reaching a peak at $\hat{n}^{club}$ and then falling.[12] Similarly, V_u^{pub} is also inverted-U shaped. The V_e^{pub} function is illustrated in Figures 2 and 3 below.

[11]The value of $\tilde{n}^{club} = n^e - n^u\left(1 - \frac{2}{\lambda^2}\right)$ can be found simply by differentiating the expression for g^{pub}. Note that for $\tilde{n}^{club} < n^e$, we need the condition $\lambda > \sqrt{2}$. Henceforth we assume this condition holds.

[12]See Appendix A.1 for derivation.

It is also useful at this stage to note the behavior of V_e^{pub} with respect to changes in parameters. An increase in λ, the effectiveness of educated individual's political action, has two effects on utility. It increases the collective political effort as it increases both the effectiveness of P^e as well the optimal choice of P^e. However, an increase in P^{e*} also increases the cost to the individual. Of these two opposing effects, the first effect dominates[13]. For every level of n^{club}, therefore, V_e^{pub} increases with an increase in λ. Further, as can be seen from the expression of $\hat{n}^{club}$, an increase in λ also shifts the peak of the V_e^{pub} function to the left. This reduces the range of n^{club} values over which the congestion effect dominates the provisioning effect. This is because greater political effectiveness leads to lesser corruption and greater public provisioning per capita.

As we are interested in the distributional aspect, n^u becomes a parameter of interest. Keeping the total population same, an increase in n^u reduces total provisioning as the political effectiveness of uneducated is lower than that of the educated. This perverse effect on g^{pub} reduces V_e^{pub} for every level of n^{club}.

An increase in the cost coefficient, δ, has the straightforward effect of increasing costs of political effort and therefore reducing every individual's optimal effort. Clearly, the reduction in cost from this change is a second order effect and its magnitude is smaller than the first order increase in δ itself, causing a fall in V_e^{pub}.

The results under public provisioning are summarized in the following proposition.

Proposition 1. *Under public provisioning, the optimal choice of political effort for educated and uneducated agents is given by (9) and (10) respectively. For both agents, this choice is decreasing in n^{pub}. The indirect utility of the educated agents, V_e^{pub}, is inverted-U shaped: initially rises with n^{club}, reaches a peak, and then falls. For every level of n^{club}, V_e^{pub} is increasing in λ and decreasing in n^u and δ.*

4.3 Club Provisioning

Now, consider the choices of an agent who opts for club provision. This agent would never expend any political effort as he derives no utility from it. We postulate that the club contribution, g_i, must cover a membership fee to join the club. This membership fee is the average fixed cost, $\frac{F}{n^{club}}$, so that the total entry fee collected from all the club members covers the fixed cost, F. The maximization problem, thus, is

[13] $n^e - n^{club} > \frac{1}{2}$ ensures that $\frac{\partial V_e^{pub}}{\partial \lambda} > 0$.

$$
\begin{aligned}
&\max_{c_i,b_i,g_i} && \alpha \log c_i + (1-\alpha)\log b_i + \tfrac{\sum_{i=1}^{n^{club}} g_i - F}{n^{club}} \\
&\text{subject to} && c_i + b_i + g_i \leq \tilde{y}_i, \\
&\text{and} && g_i \geq F/n^{club},
\end{aligned} \tag{12}
$$

where $\tilde{y}_i$ is given by (1).

The optimal consumption and bequest derived from this maximization problem are given by[14]

$$
\begin{aligned}
c_i^* =& \alpha(\tilde{y}_i - g_i^*), \\
b_i^* =& (1-\alpha)(\tilde{y}_i - g_i^*),
\end{aligned}
$$

and g_i^*, the optimal club contribution, is given by

$$
g_i^* = \begin{cases} \frac{F}{n^{club}} & \text{if } \tilde{y}_i \leq \frac{F}{n^{club}} + n^{club}, \\ \tilde{y}_i - n^{club} & \text{if } \tilde{y}_i > \frac{F}{n^{club}} + n^{club}. \end{cases} \tag{13}
$$

This result is driven by the quasilinearity of the utility function. Club contribution is (weakly) increasing in the contributor's income. Now we want to find the optimal g_i^* for an agent with a given $\tilde{y}_i$ as a function of n^{club}. Note that the function $\frac{F}{n^{club}} + n^{club}$ is non-monotonic in n^{club}. Specifically, it reaches its minimum value $2\sqrt{F}$ at the $n^{club} = \sqrt{F}$. As can be seen from Figure 1 below, a given $\tilde{y} > 2\sqrt{F}$ will intersect the $\frac{F}{n^{club}} + n^{club}$ curve at two points. Call these $n_1^{club}(\tilde{y})$ and $n_2^{club}(\tilde{y})$, with $n_1^{club}(\tilde{y}) < n_2^{club}(\tilde{y})$. Then, from (13) we have

$$
g_i^* = \begin{cases} \frac{F}{n^{club}} & \text{if } n^{club} < n_1^{club}(\tilde{y}_i) \text{ or } n^{club} > n_2^{club}(\tilde{y}_i), \\ \tilde{y}_i - n^{club} & \text{if } n_1^{club}(\tilde{y}_i) \leq n^{club} \leq n_2^{club}(\tilde{y}_i). \end{cases} \tag{14}
$$

The optimal club contribution for an individual is monotonically decreasing in the number of people who join the club.[15] This behavior, once again, is due to the free-rider effect.

As the poor never join the club (condition 5), using the optimal club contribution we can find G^{club}, the total club provision:

$$
G^{club} = \begin{cases} 0 & \text{if } n^{club} < n_1^{club}(\tilde{y}_i) \text{ or } n^{club} > n_2^{club}(\tilde{y}_i), \\ \sum_{i=1}^{n^{club}} \tilde{y}_i - (n^{club})^2 - F & \text{if } n_2^{club}(\tilde{y}_i) \leq n^{club} \leq n_2^{club}(\tilde{y}_i). \end{cases}
$$

[14]See Appendix A.1 for derivation.

[15]But, while it is linear for the interval $[n_1^{club}(\tilde{y}_i), n_2^{club}(\tilde{y}_i)]$, it is non-linear otherwise.

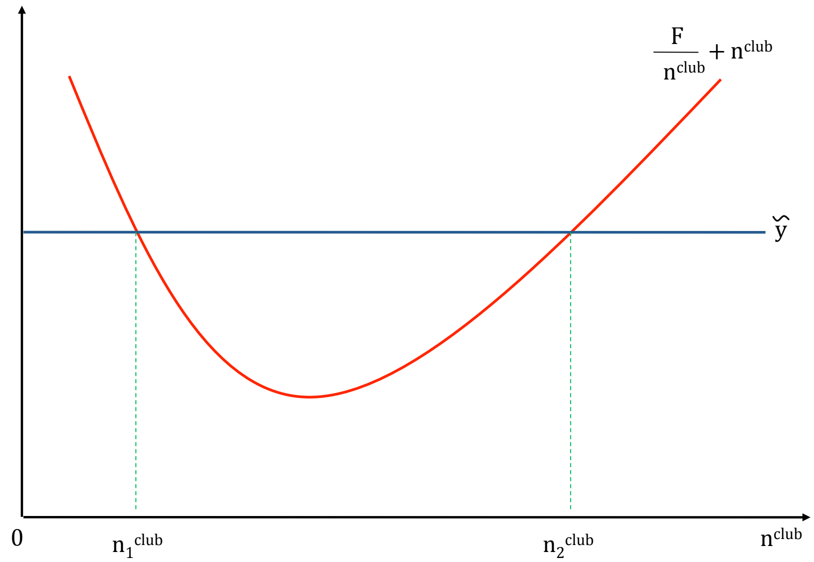

Figure 1

Just as in public provisioning, here too, we have a positive provisioning effect and a negative free rider effect as more people join the club. Unlike public provisioning, though, while the former initially dominates, the latter effect becomes more prominent for higher values of n^{club}.

Substituting (14) into (3), and given that the poor never join the club, we get the value of g^{club} as a function of n^{club}:

$$g^{club} = \begin{cases} 0 & \text{if } n^{club} < n_1^{club}(\tilde{y}_i) \text{ or } n^{club} > n_2^{club}(\tilde{y}_i), \\ \tilde{y}_i - n^{club} - \frac{F}{n^{club}} & \text{if } n_2^{club}(\tilde{y}_i) \leq n^{club} \leq n_2^{club}(\tilde{y}_i). \end{cases} \tag{15}$$

Once again, analogous to public provision, there is a congestion effect here as well. At the same time, however, there is an opposing effect due to a reduction in per capita fixed cost. The first effect eventually exceeds the second, giving g^{club} an inverted-U shape.

Plugging the optimal values of c_i and b_i, and g^{club} from (15) into (4) we get the indirect utility function of an individual who joins the club

$$V_e^{club}(x_i, n^{club}; \tau, F) = \begin{cases} A + \log\left[\tilde{y}_i - \frac{F}{n^{club}}\right] & \text{if } n^{club} < n_1^{club}(\tilde{y}_i) \text{ or } n^{club} > n_2^{club}(\tilde{y}_i), \\ A + \log n^{club} + \tilde{y}_i - n^{club} - \frac{F}{n^{club}} & \text{if } n_2^{club}(\tilde{y}_i) \le n^{club} \le n_2^{club}(\tilde{y}_i), \end{cases} \tag{16}$$

where $\tilde{y}_i$ is given by (1).

V_e^{club} is increasing for $n^{club} < n_1^{club}$ and $n^{club} > n_2^{club}$. The reason is that, in this range, as the number of agents in the club increases, the per capita fixed cost contribution required of each individual falls, while still leading to no increase in club provision. For the interim values of n^{club}, however, the curve is inverted-U shaped under the following condition[16]

$$\tilde{y}_i - 4F > 1. \tag{17}$$

In what follows, we assume condition (17) holds.[17] To understand this shape, note that g^{club} enters the utility additively. Thus the effects in play in g^{club} are also in play here. Moreover, there is further cost reduction over the entire range of n^{club} values due the free rider effect. Together, this leads to the inverted-U shape. V_e^{club} is illustrated in Figures 2 and 3 below.

It is clear from the expression of V_e^{club} that a fall in fixed cost, F, leads to an increase in V_e^{club}. This occurs because, keeping the total contribution same, the amount sunk or wasted goes down.

The following proposition summarizes the results under club provisioning.

Proposition 2. *The optimal contribution of an educated agent opting for club provisioning is given by (15). The indirect utility of such an agent, V_e^{club}, is inverted-U shaped for $n_1^{club} \le n^{club} \le n_2^{club}$ and increasing otherwise. For every level of n^{club}, V_e^{club} decreases with the fixed cost, F.*

5 Equilibrium

In this section we characterize the best response functions of individuals in order to find the equilibria. The optimal choice (between public and club provisioning) for the agents depends on the relative values of V_e^{club} and V_e^{pub}. Depending on parametric conditions, two separate possibilities may arise as illustrated in Figures 2 and 3. This result is similar in spirit to Eicher

[16] See Appendix A.1 for detailed derivation.

[17] Note that this condition is required to allow for the second possibility (Figure 3), which is the more interesting case (see Section 5). Without this condition, the V_e^{club} function may be rising or falling in the above range. In both cases we get an equilibrium that is qualitatively similar to the one in Figure 2.

et al. (2009) who also find multiple possibilities depending on parameters of the model. We first explain how the second possibility (Figure 3) gives rise to multiple equilibria. Then we discuss equilibrium selection using stochastic adaptive dynamics borrowed from evolutionary game theory.

5.1 Multiple Equilibria

First, note that for $n^{club} < n_1^{club}(\tilde{y}_i)$ and $n^{club} > n_2^{club}(\tilde{y}_i)$ the V_e^{club} function for a given agent is always below the V_e^{pub} function. This happens because, for these ranges $g^{club} = 0$, while $g^{pub} - C(P_i) > 0$. Moreover, if the individual is availing of club provision, then a further contribution of $\frac{F}{n^{club}}$ has to be made from the income, while no such contribution is required in case of public provisioning. For the interim range where both functions are inverted-U shaped, depending on whether condition (A.9) (see Appendix A.2 for details) is satisfied or not, the V_e^{club} function may or may not intersect the V_e^{pub} function. This gives rise to the two possibilities mentioned above. These are illustrated in Figures 2 and 3.

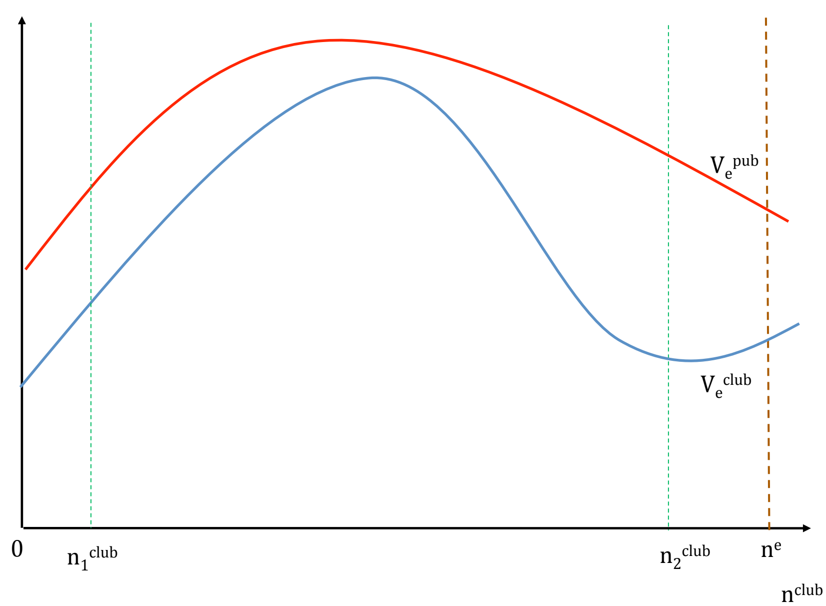

Figure 2

The *first possibility* is represented by Figure 2. In the figure, the blue curve represents the

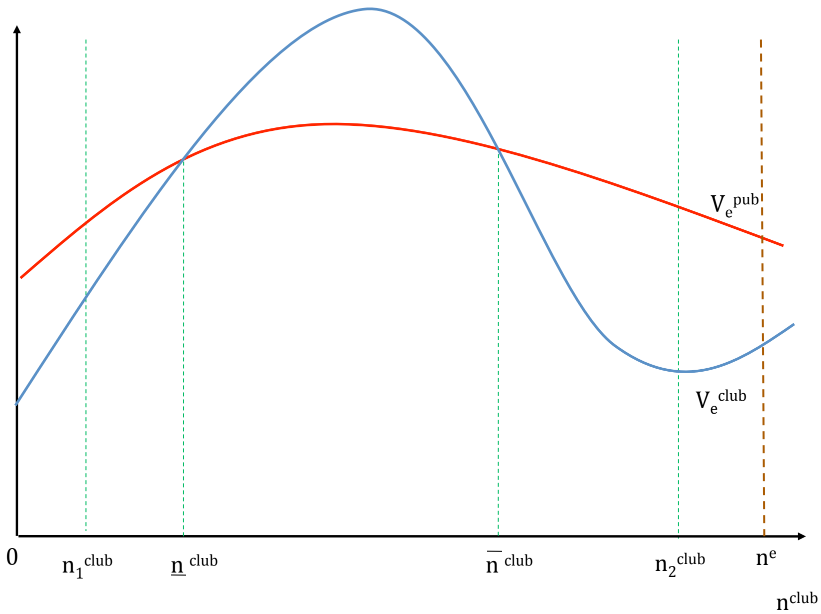

Figure 3

V_e^{club} function and the red curve represents the V_e^{pub} function. As the V_e^{pub} function always dominates the V_e^{club} function, the agent always opts for public provisioning. Therefore, the unique equilibrium occurs at $n^{club} = 0$ and everybody opts for public provisioning.

The more interesting *second possibility* is depicted in Figure 3. Here the best response of an agent is to opt for the club if the number of people in the club lies between $\underline{n}^{club}$ and $\bar{n}^{club}$ as, for these values, $V_e^{pub}(x_i, n^{club}; .) < V_e^{club}(x_i, n^{club}; .)$. For all other values of n^{club}, agents opt for public provisioning. Formally, let the agent's set of actions be $A = \{C = Club, P = Public\}$. The best response function of agent i, $B_i(n^{club})$, can be written as

$$B_i(n^{club}) = \begin{cases} P & \text{if } n^{club} \leq \underline{n}^{club}, \\ C & \text{if } \underline{n}^{club} < n^{club} \leq \bar{n}^{club}, \\ P & \text{if } \bar{n}^{club} < n^{club}, \end{cases} \tag{18}$$

where n^{club} is inclusive of agent i. This gives rise to two equilibria. In the first equilibrium, nobody opts for club provisioning. As everybody avails public provisioning, we refer to this as 'public equilibrium'. In the second, $\bar{n}^{club}$ individuals choose to join the club and the remaining $n^e - \bar{n}^{club}$ agents stay in public provisioning.[18] We refer to this equilibrium as 'club equilibrium'. The multiple equilibria arise due to a coordination problem faced by the society. The stage game as defined until now is essentially a coordination game and, depending on mutual expectations, either equilibrium can result.

To see which equilibrium entails greater corruption, we consider G^{pub}. As noted before, G^{pub} is falling in n^{club} due to the positive net provisioning effect. Therefore, any club equilibrium will feature more corruption than any public equilibrium. Thus, at least from the point of view of corruption in government, we can rank the two equilibria regardless of parametric conditions.

This is not the case with quality of public provisioning. When considering quality, g^{pub} becomes our variable of interest. Let the values of g^{pub} in the public and club equilibrium be denoted by $g_1^{pub}$and g_2^{pub} respectively. Computing these values using (2) and the fact that from (9) and (10) $P_i^{e*} = \lambda P_i^{u*}$, we get

$$
\begin{aligned}
g_1^{pub} &= \frac{(\lambda^2 n^e + n^u)P^{u*}}{N} \cdot \frac{\tau Y}{N\bar{P}\bar{\lambda}}, \\
g_2^{pub} &= \frac{(\lambda^2 (n^e - \bar{n}^{club}) + n^u)P^{u*}}{(n^e - \bar{n}^{club} + n^u)} \cdot \frac{\tau Y}{N\bar{P}\bar{\lambda}}.
\end{aligned}
$$

Comparing these values, we find that $g_1^{pub} > g_2^{pub}$ as $\lambda > 1$. The club equilibrium, since only a few educated individuals avail of public facilities (and thereby expend political effort), features lower quality of public provisioning relative to the public equilibrium. This is because education makes political action more effective ($\lambda > 1$). In view of the multiple equilibria, then, a 'coordination failure' occurs when some educated individuals opt out of public provisioning to form a club.

5.2 Equilibrium Selection

The second possibility features multiple equilibria, one of which involves lower corruption in public provisioning. Now, in order to discuss equilibrium selection,[19] we introduce stochastic adaptive dynamics. The method is borrowed from evolutionary game theory. The idea is to start from an arbitrary number of agents in the club (or, identically, in public provisioning)

[18] As all individuals are symmetric in the static model described above, the identity of who enters the club does not really matter. So, technically, while there are $\binom{n^e}{\bar{n}^{club}}$ equilibria of the second type, they are all qualitatively the same. Identity of the agents will play a role when we discuss intertemporal dynamics in section 8.

[19] In the case where there is a unique equilibrium, the issue of selection is trivial. The long run process that we describe, however, is still applicable.

and repeatedly perturb the system. This noise acts as a selection mechanism. We now describe the dynamic process with a simple noise structure, discuss possible interpretations and note a general result regarding the evolution of the system (a special case of the Markov chain tree theorem).

Consider a stage game, at stage s, that fits the features of the *second possibility*. Think of each stage as having several sub-periods. These sub-periods proceed in a discrete manner - denote a sub-period by t, where t is discrete.[20] Let the number of people who have joined the club at time t in stage s, $n_{s,t}^{club} \in K = \{0, 1, 2 \ldots, n^e\}$, denote the state of the process at that time. K, therefore is the state space. At the start of each sub-period $t+1$ an agent is randomly chosen. With a high probability, $1-\epsilon$, this agent plays the best response, $B_i(n_{s,t}^{club})$, given by (18).[21] With probability ϵ the agent randomizes between C and P (each with probability $\frac{\epsilon}{2}$). In any state, therefore, the process either stays put, moves one step to the right (if the agent plays C) or moves one step to the left (if he plays P). This gives us a simple one-dimensional Markov process. The probabilities of moving one step right and one step left are state dependent. Denote these by $R_{n_{s,t}^{club}}$ and $L_{n_{s,t}^{club}}$respectively. $F_{n_{s,t}^{club}}$ is the probability that the state remains unchanged. Their values can be computed explicitly for the best response function given in (18). The stage subscript is dropped for convenience in what follows.

$$R_{n_t^{club}} = \begin{cases} \left(1 - \frac{n_t^{club}}{n^e}\right) \frac{\epsilon}{2} & \text{if } n_t^{club} \le \underline{n}^{club}, \\ \left(1 - \frac{n_t^{club}}{n^e}\right) \left(1 - \frac{\epsilon}{2}\right) & \text{if } \underline{n}^{club} < n_t^{club} \le \bar{n}^{club}, \\ \left(1 - \frac{n_t^{club}}{n^e}\right) \frac{\epsilon}{2} & \text{if } n_t^{club} > \bar{n}^{club}. \end{cases}$$

$$L_{n_t^{club}} = \begin{cases} \frac{n_t^{club}}{n^e} \left(1 - \frac{\epsilon}{2}\right) & \text{if } n_t^{club} \le \underline{n}^{club}, \\ \left(\frac{n_t^{club}}{n^e}\right) \frac{\epsilon}{2} & \text{if } \underline{n}^{club} < n_t^{club} \le \bar{n}^{club}, \\ \frac{n_t^{club}}{n^e} \left(1 - \frac{\epsilon}{2}\right) & \text{if } n_t^{club} > \bar{n}^{club}. \end{cases}$$

$$F_{k_t} = 1 - R_{k_t} - L_{k_t}.$$

[20]This t is different from the t used in the section 8. Here it corresponds to a sub-period within a longer period, which we call a 'stage' here, denoted by s. In section 8, these subperiods are suppressed and t is used to denote the longer period.

[21]As noted earlier, the agent includes himself in assessing the current distribution.

The $n^e \times n^e$ transition matrix T describes the process.

$$T_{n^e \times n^e} = \begin{bmatrix} F_0 & R_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ L_1 & F_1 & R_1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & L_2 & F_2 & R_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & L_3 & F_3 & R_3 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & L_{n^e-1} & F_{n^e-1} & R_{n^e-1} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & L_{n^e} & F_{n^e} \end{bmatrix} \tag{19}$$

We have made two assumptions related to the random process. First, one agent is chosen at the start of each sub-period, i.e. only one agent can adjust his behavior in a given sub-period. Second, once chosen the agent with some probability may not play best response. These assumptions deliver a Markov process that is easy to work with. The extent to which our results generalize by modifying these assumption is discussed later. At this stage, it is worthwhile discussing some of the possible interpretations of this process. The standard interpretation in evolutionary theory is that agents are not completely "rational" to correctly guess the other player's choices correctly. Instead, there is a dynamic learning process. The biological interpretation is that these agents are "phenotypes" or genes that are pre-programmed to behave in certain ways. This may seem slightly at odds with our agenda, simply because we assume all agents rationally choose their effort choices and contributions. However, there are other interpretations that are related to the literature on learning in games (Fudenberg and Levine, 1998) and bounded rationality models (Rubinstein, 1998). Moreover, in the spirit of several papers on stochastic adaptive dynamics (for example, Kandori et al. (1993)) the process described here has three features. One, there is *inertia* and all agents do not instinctively react to their environment (this is related to our first assumption). Two, the agents are *myopic* in the sense that they are not forward looking and base their best response on the current state only. Three, there is always some positive probability with which the agents do not play best response. This is the *mutation* or *experimentation hypothesis.* This relates to our second assumption. There can be alternative interpretations of the two assumptions within a fully rational framework as well. That only one agent is randomly chosen at the start of each sub-period can be thought of as only one agent being 'given the opportunity' each sub-period to make a choice. This can be justified as a simplifying assumption, or with the notion that time intervals are very short. Moreover, as discussed later, this can be relaxed. An agent's randomization when given the opportunity may be seen as the agent lacking information or making a mistake. In fact, a very small ϵ places minimal restrictions on rationality, and can be interpreted as the probability with which an agent might behave irrationally. Henceforth, we use the convention of an agent being given the opportunity each period, with ϵ being the probability that he lacks information about other people's actions.

Now, we consider a result related to the long run distribution, which is a special case of the Markov chain tree theorem. This is analogous to Young (2008).[22] First we define a *k-tree*. Given any state k, a *k-tree* is a directed tree T_k consisting of all right transitions from states to the left of k and all left transitions from states to the right of k.

The Markov process that we have described will have a unique stationary distribution.[23] We however need not compute this distribution explicitly. By the following theorem, it is sufficient to only look at the order of magnitude of the probability of different states to identify the unique stochastically stable set.

Theorem 1 *(Markov Chain Tree Theorem - Special case).* For one dimensional chains, the long run probability of being in state k is proportional to the product of the probabilities on the edges of the directed tree T_k. Formally,

$$\pi_k \propto \left(\prod_{x<k} R_x\right)\left(\prod_{x>k} L_x\right).$$

Proof: See Appendix A.4.

This result allows us to compute the order-of-magnitude of the probability of each state without worrying about its exact magnitude. All we require is the directed tree for that state. To see how this works, consider the directed tree in Figure 4. The figure is drawn for a small state space to help explain the dynamics.[24]

Look at the first part of the diagram. Here the solid arrows have high probability and represent the direction of the best response. The dashed arrows arise due to randomizations and therefore have low probabilities. The convenient thing about a one-dimension process like the one depicted above is that there is a unique k-tree for each state k. As an example, the second part of the diagram shows the unique $\bar{K}$-tree. The order of magnitude of the probability of each state k is simply ϵ raised to the power that equals the number of dotted arrows coming towards k. In the diagram, the long run probability of being in state $\bar{K}$ is

[22] A more general and complete treatment can be found in Sandholm (2010).

[23] See Lemmas 1 and 2 in Appendix A.4.

[24] This explanation and depiction closely follows Young (2008). While computing the order of magnitude of the different points of the stationary distribution, we can only look at the directed tree although the system can reach the same point by following other paths.

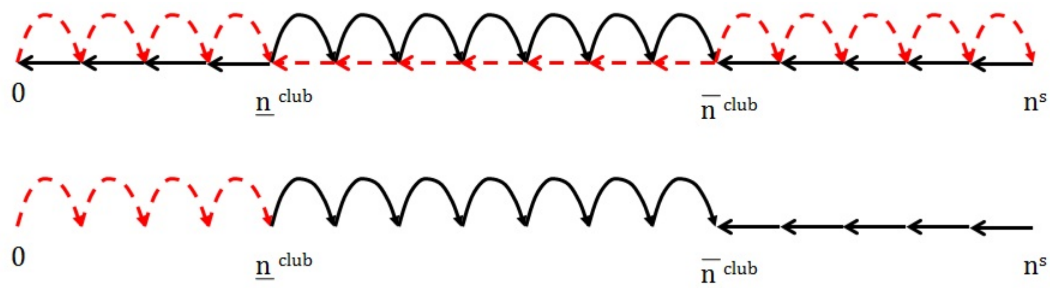

Figure 4

proportional to ϵ^4. This idea can be generally applied to all the possibilities that arise in our model.

The above diagram also shows how small but persistent shocks lead to equilibrium selection. In the above diagram, for instance, there are two stable states, namely $\bar{K}$ and 0, in a deterministic setup where all players play best response. With the perturbations added, however, the probability of being in state $\bar{K}$ is an order of magnitude $\frac{1}{\epsilon^3}$ greater than the probability of being in state 0. It follows that, as $\epsilon \rightarrow 0$, the long run distribution of the process is concentrated entirely on state $\bar{K}$ making it the only stochastically stable state. Another feature of the above dynamics is that stochastically stable state does not correspond to the Pareto optimal equilibrium but rather the risk dominant equilibrium as first pointed out by Kandori et al. (1993). A Nash equilibrium is considered risk dominant if it has the largest basin of attraction (i.e. is less risky) (Harsanyi and Selten, 1988). The basic intuition of this result can be gained from the above diagram. The "easier" it is to get to a state regardless of where the system begins, the higher is the long run probability of being in that state in the long run (or the selection bias).

A word on the specific noise process that we have assumed is in order at this stage. The assumption of one agent mutating each period allows us to work with one dimensional Markov chains which makes the analysis easier and the intuition clearer. This assumption is similar to the one made by Young (2008) in his exposition and a special case of what Samuelson (1994) calls "best response dynamics with inertia". Moreover, as discussed by Blume (2003), the result of Kandori et al. (1993) is preserved to the extent that the mutation satisfies some symmetry criteria and the payoffs observe some regularity conditions.

Now we can use Theorem 1 to examine our model in a manner similar to what is described above. Note that the the Markov process that we have described gives rise to an irreducible Markov chain so the theorem can be applied. From the dynamics corresponding to the matrix T, we can infer the following two cases.

Case 1. $\bar{n}^{club} > 2\underline{n}^{club}$. The probability of nobody going for club provisioning is larger than that of $\bar{n}^{club}$ people going for it by a factor of $\frac{1}{\epsilon^{(\bar{n}^{club}-2\underline{n}^{club})}}$. It follows that as $\epsilon \rightarrow 0$, the long run distribution of the process is concentrated entirely on state $\bar{n}^{club}$ making it the only stochastically stable state (in an manner identical to the above example).

Case 2. $\bar{n}^{club} < 2\underline{n}^{club}$. The probability of nobody going for club provisioning is smaller than that of $\bar{n}^{club}$ people going for it by a factor of $\frac{1}{\epsilon^{(\bar{n}^{club}-2\underline{n}^{club})}}$, which makes 0 the unique stochastically stable state.

The following proposition summarizes the equilibrium possibilities.

Proposition 3. *Two possibilities arise depending on parametric conditions. Let n^{club^*} denote the number of educated people who join the club in equilibrium.*
(i) *When condition* (A.9) *does not hold, there is a unique equilibrium with $n^{club^*} = 0$.*
(ii) *When condition* (A.9) *holds, there exist two equilibria. In the first, 'public equilibrium', $n^{club^*} = 0$ and in the second, 'club equilibrium', $n^{club^*} = \bar{n}^{club}$.*
Whenever there are two equilibria, the public equilibrium features lower corruption and higher quality of public provisioning. Stochastic adaptive dynamics leads to equilibrium selection, where the selection of the unique stochastically stable equilibrium depends on parametric conditions: if $\bar{n}^{club} < 2\underline{n}^{club}$ then the public equilibrium is selected, and if $\bar{n}^{club} > 2\underline{n}^{club}$ then the club equilibrium is selected.

6 Comparative Statics

In this section, we look at how the equilibrium changes with respect to parameters of the model such as effectiveness of education (λ), government repression (δ), fixed costs for setting up a private club (F) and inequality, and discuss some implications.

The comparative statics with respect to δ and λ are straightforward and intuitive. A change in these variables does not change the V_e^{club} function. A rise in λ or a fall in δ have the same the qualitative effect of shifting up the V_e^{pub} function as stated in Proposition 1. This results in a reduction of the distance between $\bar{n}^{club}$ and $\underline{n}^{club}$. Therefore, it either results in reducing the level of n^{club^*} if $\bar{n}^{club} > 2\underline{n}^{club}$ as the value of $\bar{n}^{club}$ falls, or it makes $\bar{n}^{club} < 2\underline{n}^{club}$ leading to the low corruption public equilibrium being selected.

The 'civic' content of education – awareness of rights and duties, liberal values and fostering collective norms and beliefs – is crucial to improving political outcomes. λ attempts to capture these qualitative aspects of education which have long been emphasized by political philosophers. Education has long been perceived as an instrument for affecting 'citizenship' in the political realm. In Plato's *Utopia* a holistic education in body, mind and character is viewed as the first step in the conception of an ideal polity. The role played by education in shaping beliefs and capabilities of individuals has been proposed as an explanation of the concurrent rise of mass education and nation states (Boli et al., 1985). In recent times, social theorists have construed institutionalization of education to be political indoctrination that may, under exploitative regimes, actually inhibit democracy (Bourdieu, 1986; Freire, 2000). On the other hand, Dewey (1916) is resolute in defining the overarching goal of education to be no less than fostering and maintaining a democracy. B.R. Ambedkar, Dewey's student at Columbia, was emphatically reiterating his teacher when he famously exhorted depressed castes to "educate, agitate and organize". Such diverse perspectives on the political nature of education help emphasize education as a tool for socialization. Democracy theorists have long argued for a curriculum that stresses democratic values of equality, participation and freedom as being key to shaping a democratic society. In our model, a higher value of λ represents an education system that has more pro-democratic values and illustrates the contentions of these scholars.[25] A more involved and vocal citizenry can engage in better politics, thereby putting fetters on corrupt tendencies.

Cost of political effort entails several factors. Constitutionally granted freedom of speech and freedom from government persecution are vital to a healthy democracy. Several empirical studies, such as Brunetti and Weder (2003), find a negative relation between press freedom and corruption across countries. An effective media independently reduces costs of political action by making available more accurate information and shaping public opinion. Media has also served to coordinate civil society collective action in democracies.[26] Examining panel data from Indian states, Besley and Burgess (2002) find that state governments are more responsive to falls in food production and crop flood damage via public food distribution and calamity relief expenditure where newspaper circulation is higher. Our model speaks to these, and several similar findings in the literature, as well. It is more costly, even for educated individuals, to voice discontent in repressive regimes, which impedes collective political effort. In such circumstances, it becomes feasible for potentially corrupt governments to divert more public resources for private benefits.

Another variable of interest is the fixed cost of setting up the club, F. Several policies,

[25] While we do not consider this case, it is possible to think of a negative value of λ as representing an a society with an extremely indoctrinated educated class, such as Mussolini's Italy.

[26] The Arab Spring is dramatic example of this. In another instance, in India, social as well as news media was instrumental in the spread of Anna Hazare's movement against corruption.

such as more liberal private sector laws or improved technology, may lead to the reduction of F. Such a reduction leads to greater utility from the club at all levels of club membership. As a result, club formation becomes more likely in equilibrium. Moreover, if formed, the club size would be bigger as well. This may explain the sudden rise of private provisioning in emerging economies in recent years. In India, for instance, before the reforms of 1991, entry of the private sector in institutionalized healthcare and infrastructure was severely restricted due to the 'license-quota raj', leading to high fixed costs for setting up such ventures. These costs have dramatically come down in recent times with streamlining of the business environment. This may be partly responsible for the divergence in quality of public and private provisioning in the upcoming towns. In fact, reserving the provision of certain essential services for the public sector has often been motivated by such concerns (Lülfesmann and Myers, 2011). Not allowing formation of private clubs in our model unambiguously leads to improved public provision.

Our model allows for two dimensions of inequality. One way to look at an increase in inequality is to consider a rise in the wealth of the rich, x^e. An increase in the bequest received by all agent from x^e to x^e_1 leads to a change in $\tilde{y}$ to $\tilde{y}_1$. This leads to an upward shift of the V_e^{pub} function by $\log \tilde{y}_1 - \log \tilde{y}$. In the region between n_1^{club} and n_2^{club}, which is the region we want to consider in order to look at the possibility of a club equilibrium, the V_e^{club} function shifts up by $\tilde{y}_1 - \tilde{y}$. If we are in the realm of the first possibility where there is no intersection between the curves to begin with, the change in $\tilde{y}$ might lead to an intersection as the vertical shift in V_e^{club} is greater than the vertical shift in V_e^{pub}. If we are in the second possibility, then this change makes the club equilibrium more likely as well as increases the value of $\bar{n}^{club}$, making the high corruption club equilibrium even worse.[27]

Another way to look at inequality, especially educational inequality, is to consider two distributions with the same total population but different number of uneducated people. A larger n^u, or correspondingly smaller n^e, shifts the V_e^{pub} curve down, as has been discussed earlier. This not only makes the club equilibrium more likely but also increases the value of $\bar{n}^{club}$, on the whole having a deleterious impact on the publicly provided goods or services.

Cross country studies, starting with Barro (1999) who finds primary schooling to be an important predictor of democracy, have yielded contrasting evidence. While Glaeser et al. (2004) find support for the effect of education on democracy, Acemoglu et al. (2005) contest

[27] This comparative statics is similar to Bhattacharya et al. (2012). However, our results are stronger as we show in section 8, under a simple bequest dynamics, that a society stuck in a poverty trap experiences an increase in inequality and the concomitant deterioration of public provision as an inevitable tendency of the dynamic system.

their findings and argue that the effect is driven by an omitted variable bias. In view of this controversy, Castelló-Climent (2008) finds that it is not the average level of education but the education attained by the majority of the society that is relevant for the implementation and sustainability of democracies. The contention, therefore, is that it is the distribution of education which matters. Our model explicitly allows for inequality in the distribution of education and develops a rationale for the above finding.

7 Equilibrium Ranking

Is the high corruption club equilibrium always associated with lower welfare? In this section we discuss ranking of the two equilibria in terms of welfare of different sections of the population (educated and uneducated, and among educated those who avail of club provision or public provision) and find that, surprisingly, the answer is *not* an unequivocal yes – there are situations where the high corruption club equilibrium may Pareto dominate the low corruption public equilibrium.

In a club equilibrium, the utility obtained by club members is equal to the utility obtained by educated individuals who have opted for public provisioning, that is, $V_e^{club}(x_i, \overline{n}^{club}; .) = V_e^{pub}(x_i, \overline{n}^{club}; .)$. Thus, in a club equilibrium, if a club with $\overline{n}^{club}$ members is formed, the utility of *all* the educated elite, irrespective of whether they avail of public or club provision, is given by $V_e^{pub}(x_i, \overline{n}^{club}; .)$. The utility of uneducated individuals are given by $V_u^{pub}(x_i, \overline{n}^{club}; .)$. On the other hand, in a public equilibrium, the corresponding utilities are $V_e^{pub}(x_i, 0; .)$ and $V_u^{pub}(x_i, 0; .)$ since the equilibrium club size is zero. For equilibrium ranking we need to compare $V_e^{pub}(x_i, \overline{n}^{club}; .)$ with $V_e^{pub}(x_i, 0; .)$ for the educated elite and $V_u^{pub}(x_i, \overline{n}^{club}; .)$ with $V_u^{pub}(x_i, 0; .)$ for the uneducated individuals.

Note that since the $V_e^{pub}(x_i, n^{club}; .)$ function is inverted-U shaped with respect to n^{club}, there exists a club size $\check{n}^e$ such that $V_e^{pub}(x_i, 0; .) = V_e^{pub}(x_i, \check{n}^e; .)$. As $\left.\dfrac{\partial V_e^{pub}}{\partial n^{club}}\right|_{n^{club}=0} > 0$ and $V_e^{pub}(x_i, n^{club}; .)$ reaches its unique peak at $\hat{n}^{club}$, it is clear that $V_e^{pub}(x_i, n^{club}; .) \gtreqless V_e^{pub}(x_i, 0; .)$ as long as $n^{club} \lesseqgtr \check{n}^e$. It follows that $V_e^{club}(x_i, \overline{n}^{club}; .) = V_e^{pub}(x_i, \overline{n}^{club}; .) \gtreqless V_e^{pub}(x_i, 0; .)$ as long as $\overline{n}^{club} \lesseqgtr \check{n}^e$. That is, *all* the educated elite are better off at the club equilibrium if the equilibrium club size $\overline{n}^{club}$ is smaller than the threshold club size $\check{n}^e$; otherwise they are worse off at the club equilibrium. The intuition for this result can be understood as follows. At the public equilibrium since all the educated elite access the public provision, there is a lot of congestion so that moving some educated elite out of public provision is better for all the educated elite, irrespective of whether they avail of public or club provision.

On the other hand, if too many educated elite move out of public provision then there are too few educated people to raise their voice for public provision and too many people accessing the club provision leading to a congestion there. Since the $V_u^{pub}(x_i, n^{club}; .)$ function is inverted-U shaped, by the same logic it follows that there exists a club size $\check{n}^u$ such that $V_u^{pub}(x_i, \overline{n}^{club}; .) \gtreqless V_u^{pub}(x_i, 0; .)$ as long as $\overline{n}^{club} \gtreqless \check{n}^u$. These threshold club sizes are given by

$$\begin{aligned} \check{n}^e = & \quad 2N - \frac{\lambda^2 N^2}{\lambda^2 n^e + n^u - \frac{\lambda^2}{2}}, \quad \text{for the educated,} \\ \check{n}^u = & \quad 2N - \frac{\lambda^2 N^2}{\lambda^2 n^e + n^u - \frac{1}{2}}. \quad \text{for the uneducated,} \end{aligned}$$

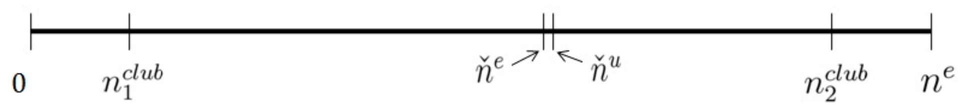

Figure 5

We can see $\check{n}^u > \check{n}^e$ and this difference is given by

$$\check{n}^u - \check{n}^e = \frac{\lambda^2 N^2(\lambda^2 - 1)}{2(\lambda^2 n^e + n^u - \frac{\lambda^2}{2})(\lambda^2 n^e + n^u - \frac{1}{2})} > 0.$$

Now we can discuss welfare comparisons between club equilibrium and public equilibrium in terms of these thresholds. As long as the equilibrium club size $\overline{n}^{club}$ falls within the range $[n_1^{club}, \check{n}^e]$, both the educated and the uneducated are better off in the club equilibrium, while if the equilibrium club size falls within $[\check{n}^u, n_2^{club}]$ they are better off in the public equilibrium. In the interim range, $(\check{n}^e, \check{n}^u)$, the educated are better off in a public equilibrium and the uneducated are better off if the educated elite forms a club. However, this range is very small relative to any reasonable population size and can be ignored for all practical purposes. Thus, $[n_1^{club}, \check{n}^e]$ can be interpreted as the 'Pareto improvement set' vis-a-vis the public equilibrium. This is because for club sizes in the range $[n_1^{club}, \check{n}^e]$, the congestion effect in public provisioning is low enough compared to the sum of net provisioning effect and cost reduction effect. Moreover, club members too are better off as compared to the public equilibrium as they enjoy sufficiently high levels of per capita private provisioning. Therefore, if the equilibrium club size remains sufficiently small, it may actually be welfare improving to form the club.

It is also worthwhile to understand the parametric configurations under which the club equilibrium Pareto dominates the public equilibrium. First, consider a relatively small value of F. Since the threshold club size $\check{n}^e$ is unaffected by F but a relatively small value of F leads to a relatively higher club size in equilibrium, the club equilibrium is more likely to be the inefficient one.[28] Therefore, it is possible that an improvement in club technology in fact leads to lower welfare, even for the club members. This happens because of the congestion externality – as the club becomes more attractive, larger club sizes may lead to overcrowding. On the other hand, relatively small values of δ or x^e, while leaving $\check{n}^e$ unchanged, has the effect of reducing the equilibrium club size. Consequently, the club equilibrium is less likely to be the inefficient one.

However, we cannot make such statements for the parameters that affect the threshold, $\check{n}^e$, itself. For instance, a relatively large value of λ results in a higher value of $\check{n}^e$, leading to shrinking of the Pareto improvement set. In other words, as the educated become more effective politically, the provisioning effect dominates the congestion effect for a larger range of n^{club} values. At the same time, however, a large value of λ reduces the equilibrium club size as well. Since, both these changes are in the same direction, the welfare effect is ambiguous. Similarly, a relatively large n^u, or correspondingly a small n^e, increases $\check{n}^e$ as well as the equilibrium club size. Here too the the net effect remains ambiguous.

8 Dynamics

In this section, we introduce a very simple bequest dynamics and study how equilibrium selection is affected over time. We begin with the poor people in the economy stuck in trap and the potential existence of a club. Specifically, we consider the second possibility. Moreover, we start by assuming that the underlying parametric conditions, especially the x^e is low enough, such that the $n^{club} = 0$ equilibrium is selected. While several other starting points are conceivable, the above configuration is used to demonstrate the basic point about how inequality in such a setup may be self-driven and eventually leads to excessive deterioration of public services.

We begin with two levels of wealth, x_0^e and x_0^u, that carry the same interpretation as above. The t subscript denotes time, which is discrete. Let C_t denote the set of people who

[28] To be sure, a small value of F leads to a lower n_1^{club}, causing a relatively larger Pareto improvement set. However, this does not affect the welfare ranking of the club equilibrium vis-a-vis public public equilibrium because the equilibrium club size moves in the opposite direction.

are in the club in period t and $C_t^{'}$ be its complement. From the individuals' optimization in section 4, we have the following bequest dynamics for dynasty i:

$$x_{t+1,i} = \begin{cases} (1-\alpha)\tilde{y}(x_{t,i}) & \text{if } i \in C_t^{'}, \\ (1-\alpha)[\tilde{y}(x_{t,i}) - g_{t,i}^*] & \text{if } i \in C_t, \end{cases} \tag{20}$$

where $\tilde{y}(x_{t,i})$ is defined in (1).

From (20), the bequest line for the poor is given by:

$$x_{t+1,i} = (1-\alpha)(1-\tau)y_u + (1-\alpha)x_{t,i}.$$

For the poor to be stuck in a trap, we impose the conditions that the steady state bequest level for the poor is below the cost of education.

$$e > \frac{(1-\tau)y_u(1-\alpha)}{\alpha}.$$

Moreover, we assume that $x_0 < e$. These two together ensure that anybody who is uneducated in period 0 stays uneducated forever. Now, we want to concentrate on the bequest dynamics of the educated class. Let n_t^{club} denote the equilibrium number of people in the club at time t. As we are in the realm of the second possibility, we know that n_t^{club} lies in the interior solution range of g_i^*. Given this knowledge, we can use (20) to write equations for the bequest lines of the educated dynasties that are in club as well as in public provisioning. For the educated dynasty i, the bequest line is given by

$$x_{t+1,i} = \begin{cases} (1-\alpha)(1-\tau)y_e + (1-\alpha)x_{t,i} & \text{if } i \in C_t^{'}, \\ (1-\alpha)n_t^{club} & \text{if } i \in C_t. \end{cases}$$

These lines are depicted in Figure 6. Note the following important features about each of these lines. First, the line for those in public provisioning is upward sloping, while for those in club is flat. This implies, in any period, everybody in the club passes the same amount of bequest, regardless of the history of their dynasty. This feature owes to the quasilinearity of the preferences assumed. Moreover, the bequest line for $i \in C_t$ is always lower than the one for $i \in C_t^{'}$ as, on top of the common wage tax, the club members need to pay the additional club contribution, $g_{t,i}^*$. Therefore, we have that starting from period 0, where (rather artificially) all educated people start with the same amount of wealth level, the ones that opt for public provisioning leave a higher level of bequest. Given the above facts, we are in a situation similar to Figure 7.

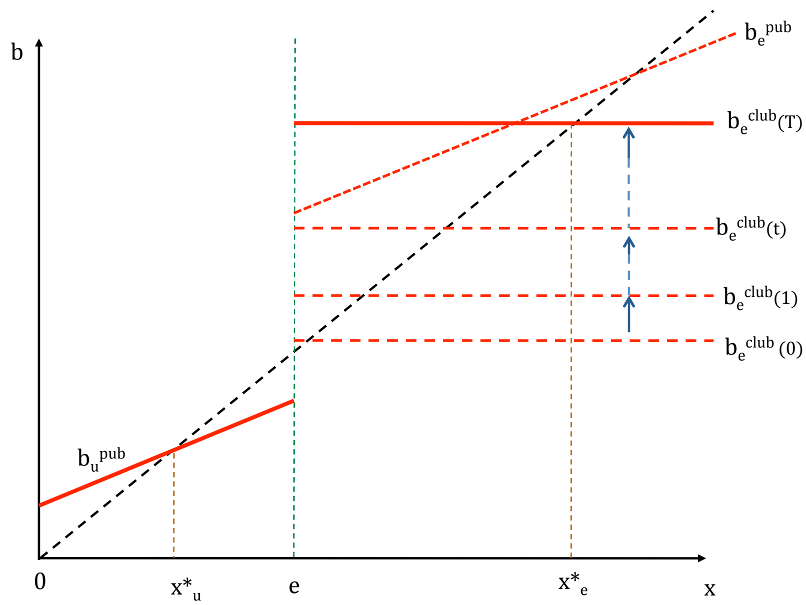

Figure 6

Starting from a homogenous educated class in period 0, we get income differentiation in period 1 as those in the club leave a lower bequest compared to those who opt for public provisioning.[29] In Figure 7, to begin with, all educated people have the same wealth level which generates income $\tilde{y}$. In the club equilibrium, n_0^{club} of these people are in the club as shown in the figure. Suppose, now, the $n^e - n_0^{club}$ people not in the club experience a greater increase in wealth than the others. For the former the income increases to $\tilde{y}_1$, and for the remaining it goes up to $\tilde{y}_2$, where $\tilde{y}_1 > \tilde{y}_2$. Call the first set of people group A, and the second set group B. Call the initial stage period 0 and the stage where there is income differentiation period 1. The dotted lines represent the position of the relevant curves in period 0 while the solid lines represent the same in period 1. There is a vertical shift of the V_e^{pub} and V_e^{club} curves of both groups from period 0 to 1. The important point to note, however, is that while V_A^e rises by more than V_B^e (the changes being $\log \tilde{y}_1 - \log \tilde{y}$ and $\log \tilde{y}_2 - \log \tilde{y}$ respectively), the V_e^{club} function shifts up by the same amount. To see the

[29]To set the described process into motion, we require that in the first period the income of the dynasties that formed the club in period 0 must go up. If this does not happen V_e^{pub} curve will shift downward instead. For an upward shift of the V_e^{pub} curve in the first period, we require that $x_0^e < (1-\alpha)n_0^{club}$.

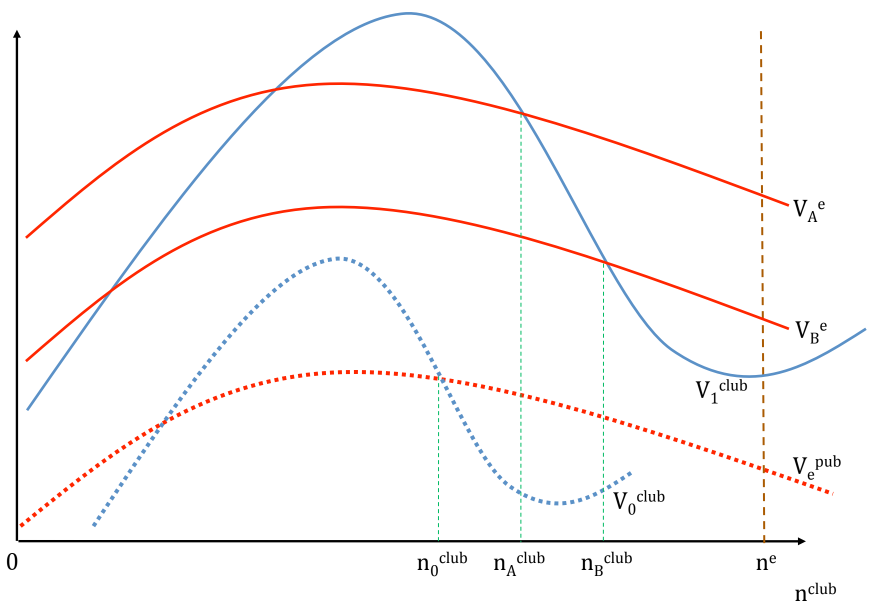

Figure 7

reason, look at the V_e^{club} function given by (16). The consumption part of the utility, given by log n^{club}, stays the same for both groups for each level of n^{club}. The part corresponding to g^{club} is the same as well because while the individuals' contributions may be different, the total contributions are pooled and evenly divided. As a result, we get shifts illustrated in Figure 7. In period 1, the club equilibrium will be at the level n_A^{club}. This is an equilibrium because beyond this level no individual from either of the groups finds it profitable to join the club. Moreover, in the new equilibrium the people who were in the club in period 0 will stay in the club even in period 1 because it is not profitable for them to leave the club given the new equilibrium configuration. The club size will increase from period 0 to period 1 and this increase will come from group A. In case, if this is the equilibrium that is not selected, this kind of a change definitely increases the probability of selection of the club equilibrium.

A situation that is qualitatively similar to the one just described occurs at every stage now. While the other line does not move, the bequest line for those in the club shifts up each period t by $(1-\alpha)(n_t^{club} - n_{t-1}^{club})$. At the same time, inequality between educated and uneducated people also increases as incomes of the educated rise. Note that as this happens, the range of n^{club} for which g_i^* has an interior solution also becomes larger. Every period features two wealth levels among the educated class. Each period more people join the club. If we began in a situation where the salient equilibrium was $n^{club} = 0$, this kind of dynamics will, in finite time, 'tip' the scales in favor of the other equilibrium. Moreover, the club size keeps increasing till every educated individual is part of the club at some time period T. The argument for everyone to join the club is the following. Suppose at any time period T there are some people who are in public provisioning. Then in the next period, those dynasties will have a higher income and therefore their indirect utility curve shifts more and some of them join the club. This cannot be a 'steady state'. Therefore, the steady state must happen when everyone has joined the club. At T, the only relevant bequest line is $b_e^{club}(T)$, which stabilizes at the level $(1-\alpha)n^e$. Moreover, once again there is only one wealth level for the entire educated class.

Perhaps the important conclusion to draw from these dynamics is that in the system described above there is a natural tendency to increase inequality. Moreover, throughout this process the quality of the public service is deteriorating. This is where improving the 'political effectiveness of education may play a vital role. Once stuck in a situation such as the second possibility, it is not possible for an economy to autonomously reduce corruption in public provision. However, as discussed in the section on comparative statics, an increase in the educational effectiveness (λ) or decreasing the cost of political effort (δ) (for instance by granting protection to whistle blowers), a society can jump to a better equilibrium.

9 Concluding Remarks

This paper studies a political economy channel through which education may impact socio-economic outcomes. Our point of departure is the postulate that education leads to more effective political action. We study the implications of this in a model of dual provisioning where people may choose between public or private (club) provisioning of certain merit goods such as healthcare and education. We analyze the role of education in people's choice of political participation and on the resultant choice between public and club provisioning. While a lower number of people opting for public provisioning is associated with lesser congestion and lesser free riding in individual political efforts, both of which lead to higher public provisioning, this effect is ultimately dominated by the fall in number of people providing political effort. This latter effect, which eventually dominates, leads to a higher level of corruption in public provisioning. Therefore, as the rich and educated set up private clubs for themselves, there are distributional consequences in provision of merit goods, owing to an increase in corruption in society under certain circumstances. Our model, therefore, provides a potential explanation for this commonly observed correlation between lower political participation of elite and poor quality of (and higher corruption in) the delivery of public services. Moreover, the possibility of private provisioning, while leading to improved provisioning for some, leads to poor quality of provisioning for the masses. This may explain why certain constitutions prohibit the entry of the private sector in provisioning of basic services.[30] In fact, our model generates two kinds of equilibria for the same parametric conditions, one of which features a private club (and therefore lower participation in public provisioning). We show that under a simple stochastic adaptive equilibrium selection mechanism, the club equilibrium becomes more likely as educational inequalities increase. This becomes relevant in light of the several empirical studies (mentioned earlier) that find educational inequality to be an important determinant of indicators of democracy and quality of governance. Moreover, we show that under a simple version of bequest dynamics this system leads to increased income inequality, which makes tipping to the club equilibrium more likely, further deteriorating the quality of public provisioning. This is where education may play a key role. Increasing the effectiveness of education (or allowing greater freedom of speech and expression) can help the economy move to a better equilibrium in terms of levels of corruption. Indeed, in our model, improving the effectiveness of education substantially can completely eliminate the possibility of club formation, despite the persistent inequality. Thus in the presence of indivisibilities in investment in education, policies which ensure free education for the poor can help the economy improve its quality of public provisioning in the long run.

The model presented above undeniably presents simplified picture. A richer model might entail a more detailed description of club formation, especially the related to collective action problems and internal workings of the club. In fact, our model completely neglects the 'industrial organization' of clubs, which is worth exploring. Also, we do not model an explicit political process or the incentives of a corrupt government. This is done primarily as a reduced form simplification in order to highlight our contention of education as being politically, and not just economically, productive. In sum, we view this model as a contribution to the incipient yet fast growing literature on the political economy linkages of education and inequality. Our main conclusions are mirrored in recent empirical findings and casual observations.

In general, this study shows the distribution of wealth is important from a macroeconomic point of view. If growth benefits only one section of society, they will ultimately arrange

[30] A similar point is made by Lülfesmann and Myers (2011).

for their own provision of public services. Moreover, this effect is permanent and can lead to the coexistence of widely different qualities of merit goods for the rich and poor. On the other hand if income growth is inclusive and can make some of the poor become educated, the inequality in per capita provisioning can be reduced.

Appendix

A.1. Individual Choices

A.1.1. Public Provisioning

Restating the maximization problem in (8), we have

$$\max_{c_i,b_i,P_i} \alpha \log c_i + (1-\alpha)\log b_i + \frac{\sum_{i\in E}\lambda P_i + \sum_{i\in U} P_i}{N\bar{P}\bar{\lambda}}.\frac{\tau Y}{n^{pub}} - \frac{\delta P_i^2}{2}$$
$$\text{s.t.} c_i + b_i \leq \tilde{y}_i.$$

The constraint will hold with equality because utility is rising in both c_i and b_i. Substituting b_i from the constraint, the first order conditions corresponding to the maximization problem for an educated person are

$$\frac{\lambda\tau Y}{N\bar{P}\bar{\lambda}n^{pub}} = \delta P^e, \tag{A.1}$$
$$\frac{\alpha}{c_i} = \frac{1-\alpha}{\tilde{y}_i - c_i}. \tag{A.2}$$

Rearranging (A.1), we get (9). From (A.2) we get $c_i^* = \alpha\tilde{y}_i$ and correspondingly $b_i^* = (1-\alpha)\tilde{y}_i$.

For an uneducated person, the choice between consumption and bequest is identical given his income. The first order condition for political effort is given by

$$\frac{\tau Y}{N\bar{P}\bar{\lambda}n^{pub}} = \delta P^u, \tag{A.3}$$

which rearranges to (10). This gives us symmetric solutions for all educated and all uneducated agents.

Now we look at the shape of the V_e^{pub} function. After simplification, we can write

$$V_e^{pub} = A + \log\left[(x_i - e) + y_e(1-\tau)\right] + \frac{\lambda}{\delta}\left(\frac{\tau Y}{N\bar{P}\bar{\lambda}}\right)^2\left(\frac{1}{n^u + n^e - n^{club}}\right)^2\left(\lambda\left(n^e - n^{club}\right) + \frac{n^u}{\lambda} - \frac{\lambda}{2}\right).$$

Differentiating this expression with respect to n^{club}, we get

$$
\begin{aligned}
\frac{\partial V_e^{pub}}{\partial n^{club}} =& \frac{\lambda}{\delta}\left(\frac{\tau Y}{N\bar{P}\bar{\lambda}}\right)^2 \left(\frac{1}{n^u + n^e - n^{club}}\right)^4 \cdot \\
& \left(-\lambda(n^u + n^e - n^{club})^2 + 2(n^u + n^e - n^{club})\left(\lambda(n^e - n^{club}) + \frac{n^u}{\lambda} - \frac{\lambda}{2}\right)\right) \\
=& \frac{\lambda}{\delta}\left(\frac{\tau Y}{N\bar{P}\bar{\lambda}}\right)^2 \left(\frac{1}{n^u + n^e - n^{club}}\right)^3 \left(-\lambda n^{club} + \lambda n^e - \lambda + n^u\left(\frac{2}{\lambda} - \lambda\right)\right).
\end{aligned}
$$

From the above, we can see that $\frac{\partial V_e^{pub}}{\partial n^{club}} \lesseqgtr 0$ iff $n^{club} \lesseqgtr \hat{n} = n^e - 1 - n^u\left(1 - \frac{2}{\lambda^2}\right)$. This implies that V_e^{pub} inverted-U shaped. Moreover, $\hat{n}$ is a global maxima.

A.1.2. Club Provisioning

Now consider the maximization problem in (12). Writing out the Lagrangian for this problem

$$
\mathcal{L} = \alpha \log c_i + (1-\alpha)\log b_i + \frac{\sum_{i=1}^{n^{club}} g_i - F}{n^{club}} + \mu_1\left(\tilde{y}_i - c_i - b_i - g_i\right) + \mu_2\left(g_i - \frac{F}{n^{club}}\right)
$$

The resulting first order conditions are

$$
\frac{\partial \mathcal{L}}{\partial c_i} = \frac{\alpha}{c_i} - \mu_1 = 0, \tag{A.4}
$$

$$
\frac{\partial \mathcal{L}}{\partial b_i} = \frac{1-\alpha}{b_i} - \mu_1 = 0, \tag{A.5}
$$

$$
\frac{\partial \mathcal{L}}{\partial g_i} = \frac{1}{n^{club}} - \mu_1 + \mu_2 = 0, \tag{A.6}
$$

$$
\left(\tilde{y}_i - c_i - b_i - g_i\right)\mu_1 \geq 0,\ \mu_1 \geq 0,\ \tilde{y}_i - c_i - b_i - g_i \geq 0 \tag{A.7}
$$

$$
\left(g_i - \frac{F}{n^{club}}\right)\mu_2 \geq 0,\ \mu_2 \geq 0,\ g_i - \frac{F}{n^{club}} \geq 0 \tag{A.8}
$$

where the last two conditions hold with complementary slackness. The constraint in (A.7) will be binding because utility is monotonically increasing in c_i. Therefore, we have two cases:

Case 1. The constraint $g_i \geq F/n^{club}$ is binding, i.e. $g_i = F/n^{club}$:

From (A.4) and (A.5), we have $\frac{\alpha}{c_i} = \frac{1-\alpha}{b_i}$ and from the budget constraint we have $b_i = \tilde{y}_i - c_i - g_i$. Putting these two together, we get the optimal values of c_i and b_i. To find the conditions under which the constraint binds, i.e., this case holds note that $\mu_2 \geq 0$. From (A.6), $\mu_2 = \mu_1 - \frac{1}{n^{club}}$. Substituting $\mu_1 = \frac{\alpha}{c_i^*}$ from (A.4), we have $\frac{1}{n^{club}} - \frac{1}{\tilde{y}_i - \frac{F}{n^{club}}} \leq 0$. This simplifies to the condition $\tilde{y}_i \leq \frac{F}{n^{club}} + n^{club}$.

Case 2. The constraint $g_i \geq F/n^{club}$ is not binding, i.e., $g_i > F/n^{club}$:

This implies that $\mu_2 = 0$. The values of c_i^* and b_i^* are found the same way as in Case 1. From (A.6), $\mu_1 = 1/n^{club}$ and from (A.4) $\mu_1 = \frac{\alpha}{c_i^*}$. Putting these together and plugging in the value of c_i^*, we have $g_i^* = \tilde{y}_i - n^{club}$. To check when this case holds, the condition $g_i^* > F/n^{club}$ implies $\tilde{y}_i > \frac{F}{n^{club}} + n^{club}$.

Now, we discuss the shape of the V_e^{club} function and derive condition (17). The sufficiency condition imposed requires the unconstrained maximizer of $A + \log n^{club} + \tilde{y}_i - n^{club} - \frac{F}{n^{club}}$, which we show to be inverted-U shaped, to be greater than n_1^{club} and less than n_2^{club}. We can find n_1^{club} and n_2^{club} by solving $\tilde{y}_i = n^{club} + \frac{F}{n^{club}}$, which gives

$$n_1^{club} = \frac{\tilde{y}_i - \sqrt{\tilde{y}_i^2 - 4F}}{2}, \quad \text{and}$$

$$n_2^{club} = \frac{\tilde{y}_i + \sqrt{\tilde{y}_i^2 - 4F}}{2}$$

On the other hand, let

$$f(n^{club}) = A + \log n^{club} + \tilde{y}_i - n^{club} - \frac{F}{n^{club}}$$

Taking derivative,

$$\frac{\partial f}{\partial n^{club}} = \frac{1}{n^{club}} + \frac{F}{(n^{club})^2} - 1$$

From here, we can see that $\frac{\partial f}{\partial n^{club}} \lesseqgtr 0$ iff $n^{club} \lesseqgtr \frac{1+\sqrt{1+4F}}{2}$, which implies f is inverted-U shaped. Moreover, the maximizer of f is given by

$$\tilde{n} = \frac{1 + \sqrt{1+4F}}{2}$$

Simplifying $n_1^{club} < \tilde{n}$, we get

$$\tilde{y}_i^2 - (4F + 1) < \tilde{y}_i\sqrt{\tilde{y}_i^2 - F} + \sqrt{1 + 4F}$$

If (17) holds, then the above inequality is satisfied. It can be checked that the same condition also guarantees $\tilde{n} < n_2^{club}$.

A.2. Parametric Conditions for the Two Possibilities

A necessary and sufficient condition for V_e^{club} to intersect V_e^{pub} is that

$$V_e^{pub}(\tilde{n}) < V_e^{club}(\tilde{n}).$$

Substituting values, this is equivalent to

$$\log \frac{\tilde{n}}{\tilde{y}_i} > (g^{pub} - C(P_i)) - g^{club},$$

which in terms of primitives of the model gives

$$\log \frac{\tilde{n}}{\tilde{y}_i} > \left[\frac{\lambda}{\delta} \left(\frac{\tau Y}{(n^u + n^e - \tilde{n}) N \bar{P} \bar{\lambda}} \right)^2 \left((\lambda \left(n^e - \tilde{n} \right) + \frac{n^u}{\lambda}) - \frac{\lambda}{2} \right) \right] - \left(\tilde{y}_i - \tilde{n} - \frac{F}{\tilde{n}} \right), \quad \text{(A.9)}$$

where $\tilde{n} = \frac{1+\sqrt{1+4F}}{2}$.

As $n_1^{club} < \tilde{n} < n_2^{club}$, we know that $\tilde{n} < \tilde{y}_i$ which makes the LHS of the above inequality negative. The RHS indicates the difference between the net benefits derived from public and private provisioning, which has to be sufficiently greater in magnitude in order to ensure that the inequality holds.

A.3. Derivation of Condition (5)

As stated above, the condition ensures that even when all the n^e educated people are in the club it is not profitable for an uneducated person to join the club. First, note that as more educated people join the club, it becomes more attractive to join as the per capita fixed costs fall and g^{pub} increases (this effect is further reinforced under our long run dynamics from section 8 where each successive generation joining the club is richer than the previous ones). Second, as the number of uneducated persons join the club, it becomes less attractive to join because their individual contributions are lower than the individual contribution of the any person from the educated class. Therefore, the condition that it is not attractive for just one educated person to not enter the club, even when all the educated people are already in the club, makes it unprofitable for more than one uneducated person to join the club as well. For an uneducated person to not join the club, we require

$$V_u^{pub} > V_u^{club}$$

to hold. This inequality becomes more difficult to satisfy with as more educated people join the club, as V_e^{club} is increasing, and V_u^{pub} is decreasing, in the number of educated people who join the club. As a sufficiency condition, we impose that it is unprofitable for the uneducated to join the club even when all educated people are in the club. This implies the following condition

$$V_u^{pub}(n^e) > V_u^{club}(n^e)$$

which is implied by

$$\log\left(1 - \frac{\frac{F}{n^e}}{\tilde{y}_u}\right) < (g^{pub}(n^e) - C(P_i^{u*}(n^e))) - g^{club}(n^e)$$

When all n^e educated people are in the club, if one educated person joins the club, the $g^{pub} - C(P_i^{u*})$ increases as the congestion in public provisioning is reduced and the g^{club} term falls for the reason mentioned above, making the inequality even tighter. Substituting values from (11) and (10), and setting $g^{club}(n^e) = \tilde{y}_u - n^e$, which is the maximum possible value of $g^{club}(n^e)$, we get the condition (5). This condition is entirely in terms of the primitives of the model. Given a value for the right hand side, we can put an upper bound on $\tilde{y}_u$.

A.4. Proof of Theorem 1

We begin with some definitions. A state j is said to be *accessible* from state i, $i \to j$, if a system that started in i has a non-zero probability of transitioning to j at some point. State i is said to *communicate* with state i, $i \leftrightarrow j$, if $i \to j$ and $j \to i$. A set of states C is a *communicating class* if every pair of states in C communicates with each other and no state in C communicates with any state not in C. A Markov chain is said to be *irreducible* if its state space is a single communicating class. State i is *recurrent* if it has a finite hitting time with probability 1. State i is *positive recurrent* if the mean recurrence time is finite. The following results are standard in Markov Chain theory[31].

Lemma 1. An irreducible Markov Chain with finite state space is positive recurrent.

Lemma 2. If a Markov Chain is positive recurrent, then a unique stationary distribution exists.

Consider an irreducible Markov chain with a finite state space $K = \{0, 1, \ldots, E\}$. It has a unique stationary distribution. Call this distribution π. Then, with probability 1, the relative frequency of being in state z equals π_z, which is independent of the initial state. Note that

$$\pi_z = \pi_{z-1}R_{z-1} + \pi_z F_z + \pi_{z+1}L_{z+1}$$

[31] See Appendix 11.A of Sandholm (2010) for statements and proofs.

where $R_z + F_z + L_z = 1 \ \forall z \geq 1$. This implies

$$\begin{array}{llcll} & \pi_{z-1}R_{z-1} + \pi_{z+1}L_{z+1} & = & \pi_z R_z + \pi_z L_z & \\ \text{or,} & \pi_{z+1}L_{z+1} - \pi_z R_z & = & \pi_z L_z - \pi_{z-1}R_{z-1} & \\ \text{or,} & \pi_z L_z - \pi_{z-1}R_{z-1} & = & \pi_{z-1}L_{z-1} - \pi_{z-2}R_{z-2} & \\ & & \vdots & & \\ \text{or,} & \pi_2 L_2 - \pi_1 R_1 & = & \pi_1 L_1 - \pi_0 R_0 & \text{(A.10)} \end{array}$$

where the last expression is due to simple recursion. Now,

$$\begin{array}{lrcll} & \pi_0 & = & \pi_1 L_1 + \pi_0 F_0 & \\ \text{or,} & \pi_0(1 - F_0) & = & \pi_1 L_1 & \\ \text{or,} & \pi_0 R_0 & = & \pi_1 L_1 & \text{as } F_0 + R_0 = 1 \\ \text{or,} & \pi_1 L_1 - \pi_0 R_0 & = & 0 & \text{(A.11)} \end{array}$$

Putting (A.10) and (A.11) together, we get the following E equations:

$$\pi_{z+1}L_{z+1} = \pi_z R_z \quad \forall E > z \geq 0 \tag{A.12}$$

This system of equations defines the stationary distribution. These are known as the detailed balance conditions. For stationarity, in the long run, a process must transit from z to $z+1$ as often as it transits from $z+1$ to z. These E equations and the constraint $\sum_{i=0}^{E} \pi_i = 1$ can together be solved to obtain the $E+1$ unknowns (π_0, π_1,. . ., π_E), the weights of the stationary distribution π. This, however, is not the route we will take. We will use the detailed balance conditions in order to prove the following theorem for one-dimensional chains, which helps us comment on the stationary distribution. We state the theorem once again and proceed to the proof.

Theorem 1 *(Markov Chain Tree Theorem - Special case).* For one dimensional chains, the long run probability of being in state k is proportional to the product of the probabilities on the edges of the directed tree T_k. Formally,

$$\pi_k \propto \left(\prod_{x<k} R_x\right)\left(\prod_{x>k} L_x\right)$$

Proof:

Let k be any particular state. The detailed balance conditions for states $x < k$ can be written as

$$\begin{aligned}\pi_1 L_1 &= \pi_0 R_0 \\ &\vdots \\ \pi_k L_k &= \pi_{k-1} R_{k-1}\end{aligned} \tag{A.13}$$

For $x = k$,

$$\pi_{k+1} L_{k+1} = \pi_k R_k \tag{A.14}$$

For $x > k$,

$$\begin{aligned}\pi_{k+2} L_{k+2} &= \pi_{k+1} R_{k+1} \\ &\vdots \\ \pi_E L_E &= \pi_{E-1} R_{E-1}\end{aligned} \tag{A.15}$$

Multiplying the RHS of all the equations in (A.13) with the LHS of (A.14), and multiplying the LHS of all the equations in (A.13) with the RHS of (A.14), we get

$$\left(\prod_{x<k} R_x\right) \pi_0 \pi_{k+1} L_{k+1} = \left(\prod_{x>k} L_x\right) \pi_k^2 R_k \tag{A.16}$$

Now, multiplying LHS of all the equations in (A.15) with the LHS of (A.16) and multiplying the RHS of all the equations in (A.15) with the RHS of (A.16), we get

$$\begin{aligned}\left(\prod_{x=0}^{k-1} R_x\right)\left(\prod_{x=k+1}^{E} L_x\right) \pi_0 \pi_E &= \left(\prod_{x=k}^{E-1} R_x\right)\left(\prod_{x=1}^{k} L_x\right) \pi_k^2 \\ \text{or, } \left(\prod_{x=0}^{k-1} R_x\right)^2 \left(\prod_{x=k+1}^{E} L_x\right)^2 \pi_0 \pi_E &= \left(\prod_{x=0}^{E-1} R_x\right)\left(\prod_{x=0}^{E} L_x\right) \pi_k^2 \\ \text{or, } \pi_k &= \sqrt{\frac{\pi_0 \pi_E}{\left(\prod_{x=0}^{E-1} R_x\right)\left(\prod_{x=0}^{E} L_x\right)}} \left(\prod_{x<k} R_x\right)\left(\prod_{x>k} L_x\right)\end{aligned}$$

Note that the term under the square root is independent of k, which gives us our result.

Q.E.D.

References

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2005): "From Education to Democracy?" *The American Economic Review*, 95, pp. 44–49.

Appadurai, A. (2004): "The capacity to aspire: Culture and the terms of recognition," *Culture and public action*, 59–84.

Barro, R. J. (1999): "Determinants of Democracy," *Journal of Political Economy*, 107, S158–S183.

Basu, A. and E. M. King (2001): "Does Education Promote Growth and Democracy? Some Evidence from East Asia and Latin America," *Washington, DC: World Bank*.

Besley, T. and R. Burgess (2002): "The political economy of government responsiveness: Theory and evidence from India," *The Quarterly Journal of Economics*, 117, 1415–1451.

Besley, T. and S. Coate (1991): "Public provision of private goods and the redistribution of income," *The American Economic Review*, 81, 979–984.

Bhattacharya, S., S. Saha, and S. Banerjee (2012): "Income Inequality, Club Formation and the Quality of Public Good: A Developing Country Perspective," *mimeo*.

Blume, L. E. (2003): "How noise matters," *Games and Economic Behavior*, 44, 251–271.

Boli, J., F. O. Ramirez, and J. W. Meyer (1985): "Explaining the origins and expansion of mass education," *Comparative education review*, 29, 145–170.

Botero, J., A. Ponce, and A. Shleifer (2012): "Education and the Quality of Government," NBER Working Papers 18119, National Bureau of Economic Research, Inc.

Bourdieu, P. (1986): "The forms of capital," *Handbook of Theory and Research of for the Sociology of Education*.

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

Brady, H. E., S. Verba, and K. L. Schlozman (1995): "Beyond Ses: A Resource Model of Political Participation," *The American Political Science Review*, 89, pp. 271–294.

Brunetti, A. and B. Weder (2003): "A free press is bad news for corruption," *Journal of Public economics*, 87, 1801–1824.

Campante, F. R. and D. Chor (2012a): "Schooling, political participation, and the economy," *Review of Economics and Statistics*, 94, 841–859.

——— (2012b): "Why Was the Arab World Poised for Revolution? Schooling, Economic Opportunities, and the Arab Spring," *Journal of Economic Perspectives*, 26, 167–88.

Castelló-Climent, A. (2008): "On the distribution of education and democracy," *Journal of Development Economics*, 87, 179–190.

Chevalier, A. and O. Doyle (2012): "Schooling and voter turnout: Is there an American exception?" Tech. rep., Discussion Paper series, Forschungsinstitut zur Zukunft der Arbeit.

Dee, T. S. (2004): "Are there civic returns to education?" *Journal of Public Economics*, 88, 1697–1720.

Delavallade, C. (2006): "Corruption and distribution of public spending in developing countries," *Journal of economics and finance*, 30, 222–239.

Dewey, J. (1916): *Democracy and Education*, Macmillan Press Limited.

Ehrlich, I. and F. T. Lui (1999): "Bureaucratic corruption and endogenous economic growth," *Journal of Political Economy*, 107, S270–S293.

Eicher, T., C. García-Peñalosa, and T. van Ypersele (2009): "Education, corruption, and the distribution of income," *Journal of Economic Growth*, 14, 205–231.

Epple, D. and R. E. Romano (1996): "Ends against the middle: Determining public service provision when there are private alternatives," *Journal of Public Economics*, 62, 297–325.

Fang, H. and P. Norman (2008): "Toward an Efficiency Rationale for the Public Provision of Private Goods," Tech. rep., National Bureau of Economic Research.

Fortunato, P. and U. Panizza (2011): "Democracy, education and the quality of government," .

Freire, P. (2000): *Pedagogy of the oppressed*, Continuum International Publishing Group.

Friedman, W., M. Kremer, E. Miguel, and R. Thornton (2011): "Education as Liberation? NBER Working Paper No. 16939." *National Bureau of Economic Research*.

Fudenberg, D. and D. K. Levine (1998): *The theory of learning in games*, vol. 2, MIT press.

Galston, W. A. (2001): "Political knowledge, political engagement, and civic education," *Annual Review of Political Science*, 4, 217–234.

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do institutions cause growth?" *Journal of economic Growth*, 9, 271–303.

Glaeser, E. L., G. A. Ponzetto, and A. Shleifer (2007): "Why does democracy need education?" *Journal of economic growth*, 12, 77–99.

Glomm, G. and B. Ravikumar (1998): "Opting out of publicly provided services: A majority voting result," *Social choice and welfare*, 15, 187–199.

Harsanyi, J. C. and R. Selten (1988): "A general theory of equilibrium selection in games," *MIT Press Books*, 1.

Helliwell, J. F. and R. D. Putnam (2007): "Education and Social Capital," *Eastern Economic Journal*, 33, 1–19.

Helsley, R. W. and W. C. Strange (1998): "Private government," *Journal of Public Economics*, 69, 281–304.

Hirschman, A. O. (1978): "Exit, Voice, and the State," *World Politics: A Quarterly Journal of International Relations*, 90–107.

Inglehart, R. and C. Welzel (2009): "How development leads to democracy: what we know about modernization," *Foreign Affairs*, 33–48.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, mutation, and long run equilibria in games," *Econometrica*, 29–56.

Lipset, S. M. (1959): "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *The American Political Science Review*, 53, pp. 69–105.

Lülfesmann, C. and G. M. Myers (2011): "Two-tier public provision: Comparing public systems," *Journal of Public Economics*, 95, 1263–1271.

Milligan, K., E. Moretti, and P. Oreopoulos (2004): "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 88, 1667–1695.

Murtin, F. and R. Wacziarg (2011): "The democratic transition," Tech. rep., National Bureau of Economic Research.

Nie, N. H. (1996): *Education and democratic citizenship in America*, University of Chicago Press.

Papaioannou, E. and G. Siourounis (2008): "Economic and social factors driving the third wave of democratization," *Journal of Comparative Economics*, 36, 365–387.

Perez-Arce, F. (2011): "The Effect of Education on Time Preferences. Working Paper." *RAND Corporation*.

Popkin, S. L. and M. A. Dimock (1999): "Political knowledge and citizen competence," *Citizen competence and democratic institutions*, 117–146.

Putnam, R. D. (1995): "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America," *PS: Political Science and Politics*, 28.

Rubinstein, A. (1998): *Modeling bounded rationality*, vol. 1, MIT press.

Samuelson, L. (1994): "Stochastic stability in games with alternative best replies," *Journal of Economic Theory*, 64, 35–65.

Sandholm, W. H. (2010): *Population games and evolutionary dynamics*, vol. 88, MIT press Cambridge.

Stepan, A., J. J. Linz, and Y. Yadav (2011): *Crafting State-Nations: India and other multinational democracies*, JHU Press.

Stiglitz, J. E. (1974): "The demand for education in public and private school systems," *Journal of Public Economics*, 3, 349–385.

Taylor, C. (1997): "The politics of recognition," *New contexts of Canadian criticism*, 98–131.

Uslaner, E. and B. Rothstein (2012): "Mass Education, State-Building and Equality," *QoG Working Paper Series*, 5, 2012.

Wantchekon, L., N. Novta, and M. Klašnja (2013): "Education and Human Capital Externalities: Evidence from Colonial Benin," Tech. rep., Princeton Working Paper.

Warren, M. E. (2011): "Voting with Your Feet: Exit-based Empowerment in Democratic Theory," *American Political Science Review*, 105, 683–701.

Wolfinger, R. E. and S. J. Rosenstone (1980): *Who votes?*, Yale University Press (New Haven).

Young, H. P. (2008): "Stochastic adaptive dynamics," *New Palgrave Dictionary of Economics, revised edition, L. Blume and S. Durlauf*, 5, 122–153.