Journal of Development Economics 87 (2008) 119–139

JOURNAL OF Development ECONOMICS

www.elsevier.com/locate/econbase

Political economy of income distribution dynamics

Lei Zhang [1]

Department of economics, Clemson University, 222 Sirrine Hall, Clemson, SC 29634, USA

Received 29 August 2006; received in revised form 8 November 2007; accepted 17 November 2007

Abstract

Income distribution varies considerably across countries; it tends to become more equal with development in some countries, but just the opposite occurs in other countries. This paper provides a theoretical investigation of the persistent differences in income distribution across countries over time. Motivated by the relationship between income distribution and public spending at different school levels for a broad range of countries over the past 30 years, the analysis centers on the role of public education where specific investments interact with political involvement by different socio-economic groups. Socio-economic groups may form lobbies to influence education policy making. The formation of lobbies is endogenous. Persistent inequality is caused by persistent lobbying efforts of the wealthy that lead to an allocation of public education spending more biased toward them.

JEL classification: D7; H4; I2; O1
Keywords: Income distribution; Allocation of public education spending; Political economy; Endogenous formation of lobby

1. Introduction

Persistent differences in income distribution across countries are a well recognized phenomenon. An existing strand of research focuses on the role of redistributive policies, in particular, public education policy. Theoretical analysis of the relationship between income distribution and total public education spending, however, is in general not consistent with empirical findings. This paper seeks to develop a theory of income distribution dynamics through the mechanism of public education policy that is consistent with empirical observations. Its point of departure is a more realistic view of public education: public education is not a single public good; on the contrary, different school levels appear to be different public goods. Public spending at different school levels exhibits distinct relationships with income distribution, and they tend to benefit different socio-economic groups. As a result, socio-economic groups have an incentive to influence the allocation of public education spending in their favor, and the resulting policies tend to preserve initial income distributions.

I formalize these ideas in a simple general equilibrium political economy model with overlapping generations, where an incumbent government determines public education spending on different socio-economic groups, and current education determines future income. The government might however be influenced by socio-economic groups. I model this political influence as lobbying by socio-economic groups in a common agency framework à la Bernheim and Whinston (1986). The uniqueness of the present model is that socio-economic groups' political influence is not assumed; rather,

E-mail address: zlei@clemson.edu.

[1] I thank my advisors Douglas Bernheim, Eric Hanushek, and Antonio Rangel for their unwavering support and invaluable advice. Helpful comments from William Dougan, David Gordon, Robert Inman, Zhi Liu, Lant Pritchett, Robert Tollison, an anonymous referee, and the Stanford Public Economics group are gratefully acknowledged.

0304-3878/$ - see front matter
doi:10.1016/j.jdeveco.2007.11.005

whether a socio-economic group will engage in lobbying is endogenously determined by its own utility maximization calculation, and is crucially dependent on current income distribution, the government's valuation of social welfare relative to political contribution, and the cost of entering the political competition. In equilibrium, different sets of lobbies may form, but the rich are invariably more likely to lobby. When different lobbies form, public education policies will be different, and over time, income distributions will evolve on different paths. In the long run, all economies may reach one unique steady state income distribution regardless of the initial distributions, or economies with different initial income distributions may reach different steady state distributions. In the latter case, we observe persistent equality or persistent inequality in different economies.

The income distribution dynamics generated by this model are in sharp contrast to that implied by the traditional median-voter model (see, for example, Meltzer and Richard, 1981; Alesina and Rodric, 1994; Persson and Tabellini, 1994). In the median-voter model, greater inequality translates into more redistributive policies; it implies that income distribution converges to one steady state in the long run. In particular, Glomm and Ravikumar (1992) and Saint-Paul and Verdier (1993) show that more unequal economy tends to support public education, and continued support of public education lowers income inequality over time. Empirical studies such as Perotti (1996) and studies reviewed in Bénabou (1996), however, find no relationship between inequality and aggregate public education spending.[2]

The present model also establishes that political influences are endogenously determined by individuals' utility maximization decision and are influenced by the economic and political parameters. Bénabou (2000) and Ferreira (2001) generate persistent inequality through less redistributive policies, but they *assume* that wealthier individuals have more political influence and that the decisive voter is richer than the median-income voter. Rodriguez (1998) shows in a lobbying model that greater inequality can lower the effective redistributive tax rates. However, he restricts the potential lobby set to the wealthier individuals, and his model is static and not suitable for studying policy and income distribution dynamics. Therefore, the present model provides a more suitable framework to analyze how changes in socio-economic environment may lead to permanent changes in income distribution dynamics.

The lobbying model in this paper serves as an instrument to formally represent various ways socio-economic groups influence policy making, in both democratic and non-democratic countries. There is anecdotal evidence that socio-economic groups lobby for favorable education policies. In the United States, the Association of Community Organization for Reform Now (ACORN), a grass-roots organization of low- and moderate-income families, lobbies persistently for better public schools in their neighborhoods; lobbying by higher education institutions is also extensive, and these lobbies tend to articulate the demands of their clients, i.e., the parents from upper socio-economic groups, and to be financed by them, for example, through alumni giving (Cook, 1998). Socio-economic groups can also influence policies by electing the "right" person or party into the office. In the US, pressures from the middle- and upper-class families have possibly been the driving force for the change in the federal college student aid programs during the 1980s and 1990s (Spencer, 1999); in Britain, middle-class influences may have helped shape the recent education policy under New Labour Party (Kerckhof et al., 1997; Thrupp, 2001). Moreover, the lack of transparency in the education budget process in many countries provides opportunities for favor-trading, affecting adversely education spending priorities and education accessibility (Hallak and Poisson, 2002). Birdsall et al. (1995) observe that "the allocation of limited fiscal resources for tertiary education, so common in Latin America, is an example of fiscal policy that reflects pressure for public spending on favored groups."

This paper is organized as follows. Section 2 provides partial correlations between income distribution and public spending at different school levels and educational attainment for a broad range of countries. These relationships motivate the theoretical analysis to follow. Section 3 sets up the political economy model and describes the equilibrium concept. Section 4 characterizes the equilibria of the stage game and discusses the income distribution dynamics. Section 5 concludes.

2. Stylized facts

This section documents three stylized facts to motivate the theoretical focus on the "composition" rather than the "level" of public education spending and on the political involvement of different socio-economic groups in the making of public education policy.

[2] Galor and Zeira (1993) in a seminal paper show that multiple steady state income distributions exist when households face liquidity constraint in their private education investment decisions. Lee and Roemer (1998) add to this model public spending on education, but the mechanism to obtain multiple equilibrium income distributions is essentially the same as Galor and Zeira.

Table 1
Summary statistics

Variable	Definition	Period	Obs	Mean	S.D.	Min	Max
Gini	Gini coefficient at beginning of decade, "accept" subset of Deininger and Squire	1970	68	0.42	0.10	0.22	0.62
		1980	69	0.40	0.09	0.21	0.63
		1990	96	0.39	0.11	0.20	0.62
		2000	106	0.40	0.09	0.24	0.63
ter_edu	Share of spending at tertiary level in total public current education spending	1970–2000	129	0.18	0.06	0.02	0.35
		1970–1990	101	0.18	0.07	0.02	0.34
		1980–2000	129	0.18	0.06	0.02	0.35
sec_edu	Share of spending at secondary level in total public current education spending	1970–2000	119	0.31	0.12	0.07	0.70
		1970–1990	94	0.29	0.11	0.07	0.61
		1980–2000	119	0.31	0.12	0.07	0.70
gdp_cap	Per capita GDP, in thousand of 1995 $US	1970–2000	134	5.19	7.88	0.11	33.67
		1970–1990	125	4.79	7.12	0.11	28.59
		1980–2000	134	5.67	8.80	0.11	39.02
ter_age	Share of tertiary-aged (18–23 years) population in total population	1970–2000	136	0.11	0.01	0.08	0.13
		1970–1990	136	0.11	0.01	0.08	0.14
		1980–2000	136	0.11	0.01	0.08	0.14
sec_age	Share of secondary-aged (12–17 years) population in total population	1970–2000	136	0.12	0.02	0.08	0.15
		1970–1990	136	0.13	0.02	0.08	0.16
		1980–2000	136	0.12	0.02	0.08	0.15
pri_age	Share of primary-aged (6–11 years) population in total population	1970–2000	136	0.14	0.03	0.08	0.18
		1970–1990	136	0.14	0.03	0.08	0.19
		1980–2000	136	0.14	0.03	0.07	0.19

Sources: Deininger and Squire (1996), World Bank (2004), UNESCO Statistical Yearbooks, and UN.
World Population Prospects Database (URL: http://esa.un.org/unpp/).

The first two stylized facts relate income distribution and public spending at primary, secondary, and tertiary schooling levels for a wide spectrum of countries over the past 30 years. They suggest that an initial income distribution can be perpetuated through public education spending at different school levels. Countries with a more unequal initial income distribution tend to spend proportionately less at secondary level and more at tertiary level, and in turn to experience a more unequal income distribution in the future.

Education spending data come from the UNESCO Institute for Statistics URL and various issues of the UNESCO Statistical Yearbook. As a measure of income distribution, the Gini coefficient is drawn from Deininger and Squire (1996) and the World Bank (2004). Income and demographics information is obtained from the World

Table 2
Coefficient estimates on Gini coefficient for three different periods

Dependent variable	1970–2000		1970–1990		1980–2000	
	OLS	Obs	OLS	Obs	OLS	Obs
(1) ter_edu	0.12^{+}[0.07]	55	0.13^{+}[0.07]	52	0.18 [0.11]	55
(2) sec_edu	−0.27* [0.13]	55	-0.27^{+}[0.14]	52	-0.34^{+}[0.20]	55
(3) ter_gdp	0.008 [0.006]	49	0.005 [0.005]	43	0.012^{+}[0.007]	51
(4) sec_gdp	−0.003 [0.008]	49	−0.0005 [0.01]	43	0.0003 [0.01]	51
(5) edu_gdp	0.01 [0.03]	58	0.007 [0.036]	56	0.049 [0.031]	58

Standard errors in the brackets; $^{+}$: significant at 10% level; *: significant at 5% level. "Obs" stands for "observations."
Note: (1) Each estimate comes from a different regression. Each regression includes per capita GDP and primary-, secondary-, and tertiary-aged population as percentage of total population. Gini coefficient is for the beginning of a period, and other variables are period averages. (2) ter_edu is share of spending at tertiary level in total public current education spending; sec_edu is share of spending at secondary level in total public current education spending; ter_gdp is share of public education spending at tertiary level relative to GDP; sec_gdp is share of public education spending at secondary level relative to GDP; edu_gdp is share of total public education spending relative to GDP.

Table 3
Coefficient estimates on education spending (dependent variable: Gini coefficient for 2000 and 1990 (t))

	Gini, t=2000 $t-1$=1970			Gini, t=1990 $t-1$=1970			Gini, t=2000 $t-1$=1980		
ter_edu	0.36**			0.44**			0.34**		
	[0.11]			[0.12]			[0.11]		
sec_edu	−0.11			−0.17*			−0.15^{+}		
	[0.09]			[0.08]			[0.08]		
ter_gdp		2.70			9.10**			3,60**	
		[1.91]			[2.29]			[1.74]	
sec_gdp		−3.39*			−6.17**			−4.52**	
		[1.71]			[1.93]			[1.21]	
edu_gdp			−0.96^{+}			0.19			−1.56**
			[0.57]			[0.46]			[0.54]
Obs	49	46	51	51	41	52	49	46	51

Standard errors in the brackets;$^{+}$: significant at 10% level; *: significant at 5% level;**: significant at 1% significance level.
Note: Each column represents a regression. Education spending is the period average from time $t-1$ to time t. Each regression also includes Gini coefficient at the beginning of a period ($t-1$) and a constant. Definitions of variables are same as Table 3.

Bank (2004) and the United Nations World Population Prospects Database URL. Public education spending is measured as the shares of spending at different school levels relative to total public education spending. Because we are interested in the income distribution evolution in the long run, and given data availability, I consider one 30-year period from 1970 to 2000 and two 20-year periods from 1970 to 1990 and from 1980 to 2000 respectively. Per capita GDP, public education spending, and shares of school-aged population are period averages; Gini coefficient is for the beginning of a period, which is taken from the year closest to the beginning of a period. This reflects the dynamic feature of the relationships. Summary statistics of the variables for each period are reported in Table 1.

First, consider how income distribution at the beginning of a period is related to public spending at different school levels over the period. Table 2 reports the coefficient estimates on the Gini coefficient in regressions of various education spending measures, controlling for per capita GDP and distribution of school-aged population. These regressions, as well as those reported in Tables 3 and 4, are intended only as illustrations of the partial correlation patterns observed in the data. Most strikingly,

Table 4

Educational attainment and income distribution
A. Proportion of population between 15 and 19 years of age that has completed at least grade 9, by wealth quintile

Decade	Quintile 1 (bottom)	Quintile 2	Quintile 3	Quintile 4	Quintile 5 (top)	# of observations
1990	0.13	0.18	0.24	0.34	0.48	44
2000	0.21	0.27	0.34	0.42	0.55	61

B. Partial correlations between Gini coefficient and Grade 9 completion rate gaps between wealth quintiles

	gap5_1	gap5_2	gap5_3	gap5_4
Gini	0.225 [0.237]	0.417 [0.198]*	0.334 [0.148]*	0.3 [0.106]**
GDP/cap	0.009 [0.017]	−0.009 [0.012]	−0.024 [0.010]*	−0.042 [0.010]**
Decade=1990	−0.022 [0.025]	−0.029 [0.022]	−0.03 [0.017]$^{+}$	−0.017 [0.013]
Constant	0.272 [0.109]*	0.149 [0.093]	0.131 [0.069]$^{+}$	0.061 [0.048]
Observations	83	83	83	83

Note: Data are from http://www.worldbank.org/research/projects/edattain/edattain.htm. Each observation is an average for a country over the 1990 or 2000 decade.
Robust standard errors clustered at country level in the brackets; $^{+}$: significant at 10% level; *: significant at 5% level; **: significant at 1% significance level.
Note: Each column in Panel B is a regression. The dependent variables are the Grade 9 completion gaps between the 5th and the 1st quintile (gap5_1), and so forth; they are decade averages (1990, 2000) for included countries. Gini coefficient is measured at the beginning of a decade, and GDP/cap is the average over the previous decade.

as shown in the first two rows, where education spending is measured as shares of spending at tertiary and secondary levels in total public education spending, countries with more unequal income distribution at the beginning of each period tend to spend proportionately more on tertiary education and less on secondary education during that period. For example, ceteris paribus, a 10-percentage point higher Gini coefficient in 1980 is associated with 3.4 percentage points less spending at the secondary level and 1.8 percentage points more spending at the tertiary level relative to total education spending in the subsequent 20-year period. This is equivalent to 11% lower spending on secondary education and 10% higher spending on tertiary education. Since spending share at primary level can be considered as a residual, these estimates suggest that higher Gini coefficient is associated with larger share spent at the primary level.

For comparison, education spending is also measured relative to GDP. When spending at tertiary and secondary levels are measured relative to GDP, the relationship is weak and inconsistent over time. Similar relationships are observed when education spending is measured relative to total government spending. These relationships are as expected. A government spends on many programs other than education, which may or may not be related to income distribution. How spending at different school levels relative to GDP, or similarly to total government spending, relates to income distribution is not clear. In row 5, the Gini coefficient does not appear to be related to aggregate education spending, echoing previous empirical findings.

Second, consider how current education spending is related to income distribution at the end of the period. Table 3 reports the partial correlations of Gini coefficients in 1990 and 2000 with education spending during the past 20- or 30-year periods, controlling for initial distribution for each period. In all three periods, when education spending is broken down by school levels, more spending at the tertiary level and less at the secondary level are associated with higher future inequality, and vice versa. However, total education spending does not appear to have a systematic relationship with future income distribution. The relationships reconfirm that total education spending may not be the policy variable most relevant for income distribution; what is more relevant is the way it is spent.[3]

The third stylized fact relates educational attainment to income distribution. Educational attainment is measured by the proportion of population between 15 and 19 years of age that has completed at least Grade 9, and is obtained from the World Bank Educational Attainment research project, which provides education outcomes for a large number of developing countries.[4] Grade 9 completion rate can indicate the potential eligibility for a college education. Panel A of Table 4 summarizes Grade 9 completion rate for different wealth quintiles. In the 1990s, almost half of the 15–19 year olds from the top quintile completed at least Grade 9, while only one-third of those from the next quintile did so, and just above 10% of those from the bottom quintile completed grade 9. The percentages improved overall in the 2000s, but the gaps did not shrink much. Panel B of Table 4 reports the partial correlations between Gini coefficient and the wealth gaps in Grade 9 completion rate. While the gap between the top and bottom quintiles is not correlated with Gini coefficient, the gaps between the top quintile and the three middle quintiles are significantly larger in countries with more unequal income distributions. Thus, not only does public spending on tertiary education benefit disproportionately the rich,[5] but this is also more so in countries with more unequal income distributions. Therefore, a division of public spending over different education levels is, to a significant extent, a division of spending over different socio-economic groups, and the division is more distinct in more unequal societies. This is key to understanding the mechanism of income distribution perpetuation through the allocation of public education spending.[6]

3. Model

This section sets up a political economy model of the allocation of public education spending on the rich, the middle class, and the poor. The model highlights the potential influences of socio-economic groups on public education spending decisions.

3.1. The economy

Consider a two-period overlapping generations model with three individuals in each generation. At the beginning of the second period of life, each individual

[3] Using other inequality measures in the regression analysis, such as the income share of the top quintile and the income share of the third and fourth quintiles, generates qualitatively similar partial correlations, but at the expense of considerable reduction in the number of observations.

[4] The URL for the project is http://www.worldbank.org/research/projects/edattain/edattain.htm. The education outcome information is derived from an analysis of various household survey data using methodology detailed in Filmer and Pritchett (1999, 2001).

[5] This is true also for developed countries, as documented by, for example, Shavit and Blossfeld (1993), Cameron and Heckman (1998).

[6] Lloyd-Ellis (2000) and Su (2004), taking education policy as exogenous, show the differential income distribution effects of public spending on basic and advanced school levels.

has one child. Let $\mathcal{I}=\{1,2,3\}$ be the set of households in each period. Each household represents a different socio-economic group: the rich, the middle class, and the poor.

Individuals go to school when young. Schools are completely publicly financed. Each child i goes to a different school with spending level e_i. The model thus assumes that education spending is perfectly targeted to each income group. As such, modelling the division of public spending over different schools is equivalent to modelling the division of public spending at different education levels. If the school attended by the poor offers only primary education, the school of the middle class both primary and second education, and the school of the rich primary, secondary, and tertiary education, then the differences in spending between the schools of the middle class and the poor and between the schools of the rich and the middle class are the spending on secondary and tertiary levels respectively. Ceteris paribus, more spending on the schools attended by the poor, the middle class, and the rich is equivalent to more spending on primary, secondary, and tertiary education levels respectively.

Schooling is the only way to accumulate human capital. As an adult in the next period, i's wage income is a concave function of e_i, $w_i^{t+1}=f(e_i)=2e_i^{1/2}$.[7] The aggregate income of the economy is $y=w_1+w_2+w_3$.[8] Let s_i be the income share of household i, then $s_i=w_i/y$ and $\sum s_i=1$. Let $s=(s_1, s_2, s_3)$.

The utility function of adult i is:

$$u_i = c_i + \delta_i f(e_i), \tag{1}$$

where c_i is consumption; δ_i measures adult i's valuation for his child's education or future income. Assuming paternalistic altruism of the parents is for simplicity and is supported by existing empirical studies (for example, Pollak, 1988; Altonji et al., 1992). I assume that, for the first generation, $\delta_i \geq \delta_j$ if $w_i \geq w_j$.[9] δ_i can also be interpreted as a child's inherited ability from his parent; then children from wealthier families have a more efficient education production function.

Differences in the initial income and the valuation of education lead to differences in the demand for education spending, which, mediated through the political process, eventually lead to differences in actual education spending on each child.

3.2. Education policy making

In each generation, an incumbent government determines the proportional income tax rate τ and how to allocate the tax revenue among the three schools by choosing $\{e_1, e_2, e_3\}$. Since an adult does not have a labor–leisure choice, the tax is nondistortionary and τ is residually determined as $\tau=(e_1+e_2+e_3)/y$. Denote the policy scheme as $p=\{e_1, e_2, e_3\}$ and $\tau=\tau(p)$. Assume that the government has to maintain a balanced budget, then the feasible policy space is defined as $\mathcal{P}=\{p=(e_1, e_2, e_3)|\tau(p)\leq 1\}$.

Government's policy making can partially be influenced by lobbies representing socio-economic groups. Lobbies exert their influence by making political contributions to the incumbent government contingent on the policy outcome.[10] The political game proceeds in two stages. In the first stage, each adult decides whether or not to enter the lobbying competition; he incurs a fixed sunk cost $F>0$ if he does. The fixed cost may represent the cost of establishing the necessary connections with government officials. This is the "entry stage". Let $\mathcal{L}\subseteq\mathcal{I}$ denote a given set of individuals that lobby. In the second stage, lobbies simultaneously and noncooperatively choose their contribution schedules; each schedule maps every education spending vector that the government might choose to a contribution level. The government sets the policy and collects from each lobby the associated contribution. This "post-entry stage" has the features of

[7] $f(e_i)$ is assumed to be strictly concave so that the first order conditions are sufficient in solving the individual utility maximization problem in Section 4. $f(e_i)$ is indeed a combination of an education production function and a wage equation. In a classical Mincerian wage equation (Mincer 1970), natural logarithm of earnings is assumed to be linear in education, and hence earnings is convex in education. In a series of surveys of studies on education production function, Hanushek (1986, 1995) suggests that education spending has a stronger positive impact on education attainment at lower level of spending, while its impact is weak at high level of spending. Depending on the shape of the education production function, $f(e_i)$ may be concave or convex. If $f(e_i)$ is convex, earnings equality or inequality will be more likely to persist in the long-run (Bourguignon 1981).

[8] The model assume no economic growth; however, this is equivalent to assuming that the technology progress increases individual productivity uniformly.

[9] Empirical evidence about time discount rates during one's lifetime suggests that the poor discount the future at a substantially higher rate than the rich (Lawrance 1991, Deaton and Paxson 1994, Atkeson and Ogaki 1996, Green et al. 1996).

[10] It is well-known that individual decision to contribute to lobbying expenditure suffers a severe collective action problem. To avoid this problem, I model each socio-economic class as a single household. Alternatively, one can assume that socio-economic classes consist of many households, but each lobby has solved the free-rider problem by some device.

the common agency model following Bernheim and Whinston (1986).[11]

An equilibrium is a set of lobbies, a set of contribution functions by the lobbies, and a policy vector that satisfy subgame perfection. More specifically, each lobby sets a contribution schedule to maximize his own utility, taking as given other lobbies' contribution schedules; the Nash equilibrium contribution schedules imply an equilibrium education spending policy that maximizes the government's objective function. Taking as given other individuals' entry decisions and the equilibrium outcome of the post-entry stage, each individual makes the entry decision to maximize his utility.

Given the fixed cost incurred when entering the lobbying game, it never pays to buy just a little political influence. Therefore, one is willing to buy into the political market only when the utility gain from lobbying is sufficiently large. This feature is indispensable in producing different equilibria for different income distributions at the entry stage.

Denote $v_i=(1-\tau(p))w_i+\delta_i f(e_i)$, then v_i is i's net utility if $i\notin\mathcal{L}$ or i's gross utility if $i\in\mathcal{L}$. Let $C_i(p)\geq 0$ denote the contribution function of $i\in\mathcal{L}$. i's net utility is,

$$u_i\begin{cases} v_i & \text{if } i\notin\mathcal{L}; \\ v_i - C_i(p) - F & \text{if } i\in\mathcal{L}. \end{cases} \tag{2}$$

The government maximizes a weighted sum of aggregate social welfare $u_1+u_2+u_3$ and political contributions. Let $V=v_1+v_2+v_3$, the government utility function boils down to:

$$W(p,C(p)) = kV + \sum_{j\in\mathcal{L}} C_j(p) - k\sum_{j\in\mathcal{L}} F \tag{3}$$

where $k>0$ is the weight the government places on the aggregate social welfare.[12]

Given the simple structure of the economy, the model can be applied to other policy settings where outcomes of a policy carry over to the next generation. Public education policy is an important example of this type of policies, and it has a profound impact on an economy's income distribution.

4. Results

Each generation plays the same political game. For each stage game, I characterize the equilibrium set of lobbies and the equilibrium spending allocation by backward induction. I then discuss the income distribution dynamics as an outcome of the political equilibrium over generations.

4.1. Allocation of public education spending

As noted before, treating $\mathcal{L}$ as given, the interaction between the various lobbies and the government has the feature of a common agency problem. I apply the results of Bernheim and Whinston (1986) and Dixit et al. (1997) to characterize the equilibrium outcome in this policy selection game.

A Subgame Perfect Nash Equilibrium (SPNE) of the post-entry game, $(\{C_j^o\}_{j\in\mathcal{L}}, p^o)$, satisfies the following three conditions.[13] First, $C_j^o(p)\in\mathcal{J}$ for all $j\in\mathcal{L}$ and $p\in\mathcal{P}$, where $\mathcal{J}$ is a set of feasible contribution functions.[14] Second, p^{o} maximizes $W(p, C^{\mathrm{o}}(p))$ on $\mathcal{P}$. This describes the best response of the government, given the optimal contribution functions. The third and essential condition prescribes the optimal behavior of every lobby. For all $m, j\in\mathcal{L}$, $(C_j^{\mathrm{o}}(p^{\mathrm{o}}), p^{\mathrm{o}})$ solves the following problem:

$$\max_{\{C_j(p)\}}(1-\tau(p))w_j - C_j(p) - F + \delta_j f(e_j), \tag{4}$$

$$\text{s.t. } W(p,(\{C_m^o(p)\}_{m\neq j}, C_j(p))) \geq \max_{\{p\in\mathcal{P}\}} W(p,(\{C_m^o(p)\}_{m\neq j}, 0)). \tag{5}$$

Consider lobby j. He takes as given contribution schedules of other lobbies when deciding his own. The constraint, Eq. (5), says that j has to provide the government at least the same payoff it can receive were j to make zero contribution. Subject to this constraint, j proposes a contribution schedule to maximize his utility, as in Eq. (4). In equilibrium, this has to be true for every lobby. Therefore, in equilibrium, each lobby contributes exactly the right amount so that the government is

[11] The common agency model was first applied to political economy of policy making by Grossman and Helpman (1994) and was considerably generalized by Dixit, Grossman, and Helpman (1997). I apply intuitions discussed in these papers to solve my model.

[12] The government maximizes $W=\frac{A}{B-A}\sum_{i=1}^{3}u_i+\frac{B}{B-A}\sum_{j\in\mathcal{L}}C_j(p)$, where $B>A>0$; i.e., the government values more highly a dollar in its own pocket than a dollar in the hands of the public. Let $k=\frac{A}{B-A}$, Eq. (3) follows immediately.

[13] As in many applications of the common agency model, using the SPNE solution concept requires a crucial assumption: the existence of perfect commitment between the lobbies and the government. Campante and Ferreira (2007) extend the common agency model to situations of imperfect commitment and analyze the efficiency implications of this extension.

[14] When individuals do not face liquidity constraint, a case assumed for the most part of the paper, $\mathcal{J}$ is $C_j^o(p)\geq 0$. With liquidity constraint, $\mathcal{J}$ is that $C_j^o(p)\geq 0$ and $F+C_i(p)<(1-\tau(p))w_i$.

indifferent between whether he contributes or not, and Eq. (5) is binding. Let

$$p^{-j} = \mathrm{argmax}_{\{p\in\mathcal{P}\}} W(p, (\{C_m^o(p)\}_{m\neq j}, 0)). \tag{6}$$

Because Eq. (5) is binding, it follows that,

$$C_j^o(p^o) = W(p^{-j}, (\{C_m^o(p^{-j})\}_{m\neq j}, 0)) - W(p^o, (\{C_m^o(p^o)\}_{m\neq j}, 0)). \tag{7}$$

By definition of p^{-j}, $C_j^o(p^o) > 0$ for all $j \in \mathcal{L}$.

The model can have multiple equilibria. In a refinement of the SPNE developed by Bernheim and Whinston (1986), each lobby offers the government a contribution function that is *truthful*, which rewards the government for every change in the policy choice exactly the amount of change in the lobby's gross-of-contribution utility, provided that the contribution both before and after the change is strictly positive. Therefore, a lobby gets the same net-of-contribution utility for all policies that induce a positive contribution from him. Thus, a competition in truthful strategies boils down to noncooperative choices of the constant $\{w_j^o\}_{j\in\mathcal{L}}$, which is the equilibrium net-of-contribution but gross-of-fixed cost utility of lobby j, and

$$w_j^o = v_j(p^o) - C_j^o(p^o), \tag{8}$$

where $C_j^o(p^o)$ is given by Eq. (7). For the rest of the paper, I focus on the Truthful Nash Equilibrium (TNE).[15]

Focusing on the TNE provides a straightforward way of calculating the equilibrium policy scheme. Since Eq. (5) is always binding in equilibrium, we can substitute it into Eq. (4) for $C_j(p)$ and replace $C_m^o(p)$ with its truthful form; then the government agent's constrained maximization problem over the policy vector is,[16]

$$\max_{\{p\}} kV + \sum_{j\in\mathcal{L}} v_j. \tag{9}$$

Thus, in a TNE, the government reaches the optimal policy scheme by maximizing a weighted sum of the aggregate social welfare and the aggregate welfare of all the lobbies. This is equivalent to maximizing a weighted sum of the welfare of all the individuals, with a larger weight placed on the welfare of the lobbies. The equilibrium policy in a TNE is efficient in that it maximizes the joint welfare of the government and the lobbies.[17] Moreover, since the government also cares about the welfare of the non-lobbies, their welfare will not be driven down to a minimum.

I characterize the optimal policy scheme prescribed by Eq. (9) for four generic cases: no-lobby, one-lobby, two-lobby, and three-lobby, denoted as $\mathcal{L}=\emptyset$, $\mathcal{L}=\{1\}$, $\mathcal{L}=\{1, 2\}$, and $\mathcal{L}=\{1, 2, 3\}$. The results are summarized in Table 5. Note that it is possible that $\mathcal{L}=\{2\}$ or $\{3\}$ in the one-lobby case, and $\mathcal{L}=\{1, 3\}$ or $\{2, 3\}$ in the two-lobby case, but the equilibrium results are similar to those for $\mathcal{L}=\{1\}$ and $\mathcal{L}=\{1, 2\}$.

The benchmark case is the no-lobby case, where the government maximizes the aggregate social welfare.[18] In this case, a child receives a share of the total education spending in accordance with his parent's preference for his education, and the education policy favors no one in particular in excess of his own preference. The next generation income distribution is determined accordingly. In the one-lobby and two-lobby cases, the government attaches a higher weight to the utility of the lobby.[19] Compared to the

[15] Bernheim and Whinston (1986) prove that truthful strategies and TNE have the following desirable properties. First, the best response set of lobby j to any set of contribution functions by his opponents contains a truthful contribution function. Thus, a lobby bears essentially no cost from playing truthful strategies. Second, the TNE always results in an efficient policy choice in that it maximizes the joint payoff to the government and the lobbies. Third, the set of TNE coincides with the set of coalition-proof Nash Equilibria which are stable when non-binding communication among lobbies is possible.

[16] Because individual utility function is separable in policy and contribution, we can separately characterize the optimal policy and the net-of-contribution utilities. In the general setup of Dixit et al. (1997), optimal policies and net utilities are determined simultaneously. In the next subsection, I will turn to Eqs. (7) and (8) and characterize lobbies' net utilities.

[17] Campante and Ferreira (2007) show in a common agency framework that the efficiency property no longer holds with imperfect commitment between government and lobbies. Lobbying in general leads to an efficiency loss. Changes in income distribution have ambiguous effects on policy outcomes and the magnitude of efficiency loss. Esteban and Ray (2006), in a very different model of lobbying, obtain similar conclusions.

[18] Optimal policy for the benchmark case is obtained by maximizing the following objective function:

$$\begin{aligned} kV &= k(v_1 + v_2 + v_3) \\ &= k\sum_{i=1}^{3}\left((1-\tau)w_1 + 2\delta_i e_i^{1/2}\right) \\ &= k\sum_{i=1}^{3}\left((y-(e_1+e_2+e_3))s_i + 2\delta_i e_i^{1/2}\right) \\ &= k(y-(e_1+e_2+e_3)) + k\sum_{i=1}^{3} 2\delta_i e_i^{1/2}. \end{aligned}$$

The first order condition is $-1+\delta_i e_i^{-1/2}=0$ for $i \in \{1, 2, 3\}$. It follows that $e_i=\delta_i^2$ for i Î {1, 2, 3}.

[19] For the cases $\mathcal{L}=\{1\}$ and $L=\{1, 2\}$, the equilibrium policy is obtained by maximizing

$$kV + v_i = k(k+s_1)(y-(e_1+e_2+e_3)) + 2(k+1)\delta_1 e_1^{1/2} + 2\left(\delta_2 e_2^{1/2} + \delta_3 e_3^{1/2}\right),$$

and

$$kV + v_i + v_2 = k(k+s_1+s_2)(y-(e_1+e_2+e_3)) + 2(k+1)\left(\delta_1 e_1^{1/2} + \delta_2 e_2^{1/2}\right) + \delta_3 e_3^{1/2}$$

respectively. The first order conditions and optimal policy scheme follow immediately.

Table 5
Equilibrium policy and next generation income distribution

Lobby Case	Government objective function	Equilibrium policy, $p=\{e_1, e_2, e_3\}$	Next generation income distribution $s'=\{s'_1, s'_2, s'_3\}$
$\mathcal{L}=\emptyset$	kV	δ_1^2	$\frac{\delta_1}{\delta_1+\delta_2+\delta_3}$
		δ_2^2	$\frac{\delta_2}{\delta_1+\delta_2+\delta_3}$
		δ_3^2	$\frac{\delta_3}{\delta_1+\delta_2+\delta_3}$
$\mathcal{L}=\{1\}$	$kV+v_1$	$\frac{\delta_1^2(k+1)^2}{(k+s_1)^2}$	$\frac{\delta_1(k+1)}{\delta_1(k+1)+\delta_2 k+\delta_3 k}$
		$\frac{\delta_2^2 k^2}{(k+s_1)^2}$	$\frac{\delta_2 k}{\delta_1(k+1)+\delta_2 k+\delta_3 k}$
		$\frac{\delta_3^2 k^2}{(k+s_1)^2}$	$\frac{\delta_3 k}{\delta_1(k+1)+\delta_2 k+\delta_3 k}$
$\mathcal{L}=\{1, 2\}$	$kV+v_1+v_2$	$\frac{\delta_1^2(k+1)^2}{(k+s_1+s_2)^2}$	$\frac{\delta_1(k+1)}{\delta_1(k+1)+\delta_2(k+1)+\delta_3 k}$
		$\frac{\delta_2^2(k+1)^2}{(k+s_1+s_2)^2}$	$\frac{\delta_2(k+1)}{\delta_1(k+1)+\delta_2(k+1)+\delta_3 k}$
		$\frac{\delta_3^2 k^2}{(k+s_1+s_2)^2}$	$\frac{\delta_3 k}{\delta_1(k+1)+\delta_2(k+1)+\delta_3 k}$
$\mathcal{L}=\{1, 2, 3\}$	$(k+1)V$	Same as $\mathcal{L}=\emptyset$	Same as $\mathcal{L}=\emptyset$

Note: The lobby cases are the no-lobby, one-lobby, two-lobby, and three-lobby cases. Government maximizes a weighted sum of social welfare function and the lobbies' utilities. k is government's valuation for the social welfare; δ_i is individual i's valuation for his child's education; s_i and s'_i are the income share of individual i in current period and that of his child in the next period. Ceteris paribus, an individual obtains more education spending in the current period and his child a larger share of income in the next period when he is a lobby than when he is not.

benchmark case, the child of the lobby receives a larger share, as well as a larger amount, of the public education spending; consequently, in the next period, he gets a larger share of the total income. For a non-lobby's child, the opposite holds. In addition, with more people lobbying and more political competition, the increase in spending on a lobby diminishes and the decrease in spending on a non-lobby enlarges.

When all three individuals lobby, the government maximizes $kV+v_1+v_2+v_3=(k+1)V$. The education policy is identical to that in the no-lobby case; so is the income distribution of the next generation. Therefore, political competition has an impact on the equilibrium education policy and the future income distribution only when a subset of the individuals enter the lobbying. Compared to the benchmark case, however, every individual is worse off in the three-lobby case due to the fixed cost and positive contribution payment associated with lobbying.

When a subset of the individuals organize into lobbies, the government's preference for the aggregate social welfare, k, plays an important role in determining the relative amount spent on each child and the future income distribution. The higher the k, the more concerned the government is about social welfare, and the less biased the policy is toward the lobbies. At the limit, when k approaches infinity, the allocation of education spending, as well as the income distribution of the next generation, is identical to that in the no-lobby case, regardless of which lobbies form. Indeed, as discussed in the next section, when k is very high, individuals have little incentive to enter the lobbying game in the first place.

4.2. Endogenous formation of lobbies

Based on the equilibrium analysis of the post-entry phase, I now consider individuals' decision of whether to enter the lobbying game. Individuals make their entry decisions based on their own utility maximization calculation. Different sets of lobbies may form at the entry stage; the entry equilibrium is closely related to an economy's income distribution.

For the rest of the paper, I define the income distribution space of the first generation as $S^3=\{s|s_1\geq s_2\geq s_3;\ s_1+s_2+s_3=1\}$; thus, individual 1 represents the rich, 2 the middle class, and 3 the poor. In addition, let $D_{12}=\delta_1^2/\delta_2^2$ and $D_{23}=\delta_2^2/\delta_3^2$; $D_{12}>1$ and $D_{23}>1$ measure the differences in individuals' taste for education.

An individual's choice of whether or not to enter the lobbying game depends on the utility he would obtain from lobbying in the post-entry stage as analyzed in the previous subsection relative to the utility he would obtain if he were not to enter at all. Consider individual j. Let $G_j(\mathcal{L})$ denote the difference in j's net utilities when $j\in\mathcal{L}$ and when $j\notin\mathcal{L}$, taking as given the entry decision of all $i\in\mathcal{I}$ and $i\neq j$, then by Eq. (8)

$$G_j(\mathcal{L}) = w_j^o - F - v_j\left(p^{-j}\right), \tag{10}$$

where p^{-j} is identical to the policy scheme defined in Eq. (6). The first two terms determine j's net utility after paying the equilibrium contribution and the fixed cost when he lobbies; the last term denotes j's utility when he does not enter the lobbying game at all. j enters the lobbying game if and only if $G_j\geq 0$. Therefore, $\mathcal{L}$ is an equilibrium set of lobbies when $G_j(\mathcal{L})\geq 0$ for all $j\in\mathcal{L}$ and $G_j(\mathcal{L})<0$ for all $j\notin\mathcal{L}$.

To enter the lobbying game, individuals have to first incur a fixed cost F. This fixed cost essentially serves as a commitment mechanism: by not paying the fixed cost, individuals forfeit the potential to influence the policy in their favor. Eq. (10) clearly demonstrates that individuals will enter only if the gains in utility are large enough to outweigh the upfront fixed cost payment. Therefore, the magnitude of the fixed entry cost plays a critical role and provides a natural gauge to determine an individual's entry decision.

Characterization of $v_j(p^{-j})$ is straightforward: one simply substitutes p in individual utility function with one of the optimal policy schemes in Table 5. Characterizing the net utility for a lobby is more complicated because his net-of-contribution utility w_j^o is determined simultaneously with that of other lobbies in equilibrium. In addition, since the social welfare function is additive in individual utility, which is quasi-linear in consumption, government is indifferent regarding the amount of contribution by each lobby as long as the total contribution is the same. Therefore, in equilibrium, the net utilities of lobbies are indeterminate except for the case of a single lobby.

In Appendix 1, I derive the upper bound of w_j^o for each possible lobby case using Theorem 2 of Bernheim and Whinston (1986).[20] It provides the exact utility value for the one-lobby cases and for the two-lobby case with the rich and the middle class being the lobbies. In Appendix 2, these upper bounds are substituted into Eq. (10) for each individual to characterize conditions for each of the eight possible entry equilibria. This practice is harmless in proving the key results below for two reasons. First, since entry equilibria are characterized by inequalities, using these upper bounds frequently imposes more stringent conditions for most entry equilibria. Second, as discussed in Appendix 1, it can be shown that with loose restrictions on parameter values (δ_i's and k), these upper bounds provide the exact net utility values. Discussion in the remainder of this section is based on the utility gain function $G_j(\mathcal{L})$ derived in Appendix 2.

The expressions of $G_j(\mathcal{L})$ for all $j\in\mathcal{I}$ and all possible $\mathcal{L}$ are functions of model parameters (k, F, and δ's) and the initial income distribution. Income levels are irrelevant; this is because of the quasi-linearity of both individual utility function and social welfare function and that there is no liquidity constraint.

Given that an individual has only two choices at the entry stage, entering and not entering, a mixed strategy equilibrium always exists. The more interesting equilibria, however, are pure strategy equilibria. Key results of the analysis are as follows; it establishes the existence and the uniqueness of a pure strategy equilibrium at the entry stage.

When government's valuation of social welfare (k) is sufficiently large, and when the differences in individuals' taste for education, D_{12} and D_{23}, are in an intermediate range, for a wide range of entry cost, a pure strategy equilibrium exists for any initial income distribution. When an equilibrium exists, it is the unique equilibrium for most initial income distributions in the income distribution space. The unique entry equilibrium varies with the entry cost:

1. *When the entry cost is sufficiently low, for any initial income distribution, all three individuals enter ($\mathcal{L}=\{1, 2, 3\}$) is the unique equilibrium.*
2. *When the entry cost increases, for any initial income distribution, the number of lobbies decreases. When a subset of individuals lobby, the unique equilibrium is either the rich and the middle class enter ($\mathcal{L}=\{1, 2, 3\}$) or only the rich enter ($\mathcal{L}=\{1\}$).*
3. *When the entry cost is sufficiently high, for any initial income distribution, no one enters ($\mathcal{L}=\emptyset$) is the unique equilibrium.*

I prove these results in Appendix 3 by proving Lemmas A1–A4. The intuition of these results comes from a close examination of the benefit and cost of

[20] The upper bound can also be derived by directly applying the SPNE concept in the previous subsection. I illustrate this for the one-lobby case in Appendix 1.

lobbying. A lobby derives benefits from two sources. First, when he enters lobbying, he lobbies to increase public education spending on his own child. Second, he also lobbies to reduce the spending on other people's children; this will contribute to a reduction in the tax burden. Because the rich have a higher valuation for their children's education, they benefit more from the first channel; because the rich have higher income, they also benefit more from any reduction in tax rate. Adding the cost of lobbying – the contribution payment – will reduce the size of these benefits but will not change the basic property. Starting from the benchmark case of no-lobby. If the fixed cost becomes lower, benefit of the rich from lobbying is most likely to exceed the fixed cost. Therefore, $\mathcal{L}=\{1\}$ is the most likely one-lobby case. Given that the rich lobby, if the fixed cost drops further, benefit of the middle class is more likely than that of the poor to exceed the fixed cost. Therefore, $\mathcal{L}=\{1, 2\}$ is the most likely two-lobby case. When the fixed cost is very low, benefit of all three individuals from lobbying will be larger than the fixed cost, and they will all enter.

Lemmas A1–A4 establish sufficient conditions for the existence of a pure strategy equilibrium and for $\mathcal{L}=\{1\}$ and $\mathcal{L}=\{1, 2\}$ to be the unique one-lobby and two-lobby cases respectively. The sufficient conditions require D_{12} and D_{23} to be in an intermediate range.[21] The reason for maintaining these restrictions is that each individual's utility gain from lobbying increases with every individual's valuation for his child's education. Consider individual j. Intuitively, given others' strategies, a higher δ_j translates into a larger utility gain from lobbying for j since his child receives more education spending hence higher future income. In addition, as j comes to lobby, education spending on other individuals' children drops. This drop contributes to a reduction in the tax rate; the higher δ_i, the larger the reduction in the tax rate. Thus, other people's greater taste for education provides an indirect incentive for one to enter lobbying. Restrictions on D_{12} and D_{23} widen the gap in the taste for education between each pair of individuals; this essentially gives the wealthier groups more incentive to enter lobbying: the gain in future income is sufficiently large. Meanwhile, these restrictions do not widen the gap too much and essentially restrict further the incentive of the less wealthy groups to enter lobbying: the tax reduction due to the drop in education spending on others as they come to lobby is moderate, which contributes to a moderate utility gain for them.

The entry equilibrium is established by fixing the government's valuation of social welfare, k, and allowing the entry cost to vary. It can be similarly established by fixing the entry cost and letting government's valuation of social welfare vary. When the government has a higher valuation of social welfare, every socio-economic group will have less incentive to enter the lobbying competition.

The proof of Lemmas A1–A4 provides a complete partition of the income distribution space by types of entry equilibrium for a wide range of entry cost. Fig. 1 reproduces detailed results of the lemmas and illustrates typical entry equilibria for given values of individual and government preferences (k and δ's) while the entry cost increases. In each panel, the horizontal axis measures the initial income share of the middle class, $s_2 \in [0, 1/2]$, and the vertical axis that of the rich, $s_1 \in [1/3, 1]$. The triangle OED represents the income distribution space S^3. Along line OE, $s_1 = s_2$; along line DE, $s_2 = s_3$. The income share of the poor is zero along line OD, is constant and positive along any line segment parallel to OD, and reaches its highest possible level, 1/3, at point E.

First, as entry cost increases over panels (a) through (f), fewer individuals enter lobbying, but the rich are invariably more likely to enter. At very low entry cost, all three individuals enter the lobbying irrespective of their income shares, as shown in panel (a). When the entry cost increases, as shown in panel (b), the rich and the middle class still enter, irrespective of their income, but the poor's entry decision depends on his income. When the entry cost increases further, the rich still enter regardless of his income, but the middle class's entry decision depends on his income, while the poor never enter, as shown in panels (c) and (d). Panels (e) and (f) show cases with even higher entry cost. In both cases, the rich's entry decision is a function of his income, while the middle class and the poor never enter.[22]

Second, individuals' entry decision depends on their income; the entrants in general have sufficiently high income (within their respective feasible income range). This is because given fixed δ's, the benefit of lobbying simply depends on one's income. The higher the income, the more benefit a lobby obtains from a reduction in tax rate.[23] This can be seen from panels (b) through (f).

[21] These restrictions can be relaxed substantially if we ignore extreme income distributions in the proofs such as $\{s_1, s_2, s_3\} = \{1/3, 1/3, 1/3\}$ or $\{1, 0, 0\}$.

[22] The two cases, however, imply different income distribution dynamics, as discussed in the next section.

[23] At relatively low income, one may benefit from a big increase in spending on his child and small change in tax burden because others will bear a large proportion of the increased tax burden; therefore he will enter e.g., the middle class in case (c). However, this is because δ is fixed for each income group. Once we allow δ to vary more flexibly with income per se, this will not be the case.

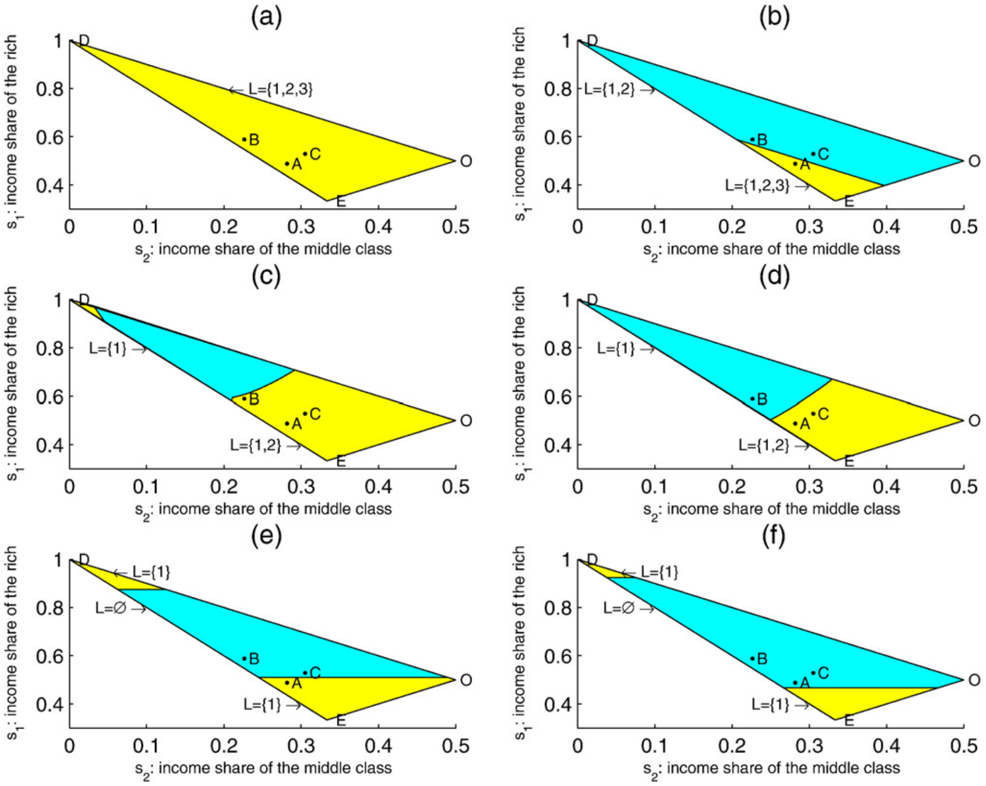

Fig 1. Typical Equilibria at the Entry Stage. Note: The valuations for children's education are $\delta_1^2=0.45$, $\delta_2^2=0.15$, and $\delta_3^2=0.1$ for the rich, the middle class, and the poor respectively. Government's valuation for social welfare is $k=2$. The entry cost increases from panel (a) to panel (f). Point A is the equilibrium next generation income distribution when $\mathcal{L}=\emptyset$ or $\mathcal{L}=\{1, 2, 3\}$; Point B is the equilibrium next generation income distribution when $\mathcal{L}=\{1\}$; Point C is the equilibrium next generation income distribution when $\mathcal{L}=\{1, 2\}$.

Third and as a direct outcome of this "income effect", focusing on equilibria where a subset of individuals enters the lobbying game and ignoring extreme income distributions, we observe that at a given entry cost, more individuals enter in economies with a more equal income distribution, and fewer individuals enter in economies with a more unequal income distribution. In panel (b), only the rich and the middle class enter in economies with a more unequal income distribution (in the dark-shaded area), while all three individuals enter when the distribution is more equal (in the light-shaded area). Similarly, in panels (c) and (d), the rich enter when an economy's distribution is more unequal (in the dark-shaded area), while both the rich and the middle class enter when the distribution is more equal (in the light-shaded area). Combined with the equilibrium education spending policies derived in Section 4.1, this result is crucial in obtaining persistent equality or inequality over time as discussed in the next section.

Also depicted in each panel are next generation's income distributions associated with different entry equilibria solved in Section 4.1; they depend on government's and individuals' preferences, and they depend on the initial income shares only through the entry decisions. Point *A* represents the equilibrium income distribution of the next generation for no-lobby or three-lobby cases; Points *B* and *C* are the respective equilibrium income distributions when the rich lobby and when both the rich and the middle class lobby. The associated Gini coefficient is the highest for Point B and is the lowest for Point A; this ordering does not depend on the numerical values of the parameters.

4.3. Income distribution dynamics

This section answers the following questions: For any initial income distribution $s \in S^3$, how does it evolve over time? Is there a steady state income distribution in the long run? The answers follow from the outcome of the stage game and crucially hinge on the values of two institutional parameters: government's preference for social welfare (k) and the entry cost of lobbying (F).

To recapitulate, given income distribution, values of these two parameters determine current generation's lobbying decision and hence the education policy. Government's preference for social welfare (k) also directly enter the calculation of current education spending. Once the current education spending is determined, next generation's income distribution, and whether the next generation in turn will act to influence the policy making, is completely determined. Small changes in the values of the government's preference and entry cost may change the path of the income distribution dynamics and the long run steady state income distributions.

Depending on the values of the government's valuation of social welfare and the entry cost, income distribution dynamics fall into one of the three broad categories. First, all initial income distributions converge to a unique steady state in the long run. This convergence occurs when the entry cost is sufficiently low or high, such that for any value of the government preference, $\forall s \in S^3$, the only entry stage equilibrium is either the three-lobby equilibrium or the no-lobby equilibrium, leading to a unique policy and next generation distribution. Facing the same parameters, the next generation will make the same entry decision, and so forth. In this case, the steady state income distribution is reached in one period. Convergence also occurs when the government's valuation for social welfare is sufficiently high but the entry cost is in an intermediate range. In this case, different entry equilibria may form for different initial income distributions, but the income distributions of the next generation are close enough because the policy is less biased towards the lobbies, and the next generation will make identical entry decision regardless of its income distributions. The unique steady state distribution is reached in two periods.

Second, different initial income distributions converge to different steady states in the long run. When the entry cost is in an intermediate range, and the government's valuation of the social welfare is not too high, different entry equilibria will form for different initial income distributions, leading to different education policies and next generation income distributions. Moreover, the next generation income distributions are sufficiently different, leading the next generation to make different entry decisions. In particular, if, over time, children make the same entry decisions as their parents, then the short-run equilibrium income distributions will be perpetuated, and economies will reach different steady state income distributions in the long run.

Third, income distribution does not converge in the long run. For somewhat different parameter values from those in the second scenario, if children make different entry decisions from their parents, then income distribution will not converge over time. The discrepancy between children's and parents' entry decision could be an outcome of the paternalistic preferences.

Fig. 1 also illustrates different income distribution dynamics. In panel (a), with one entry equilibrium for all income distributions, different economies all converge to the unique long run distribution, Point A, in one period; in panels (c) and (f), different entry equilibria form for different initial income distributions, but all economies converge to the unique long run distribution, Point C and Point A respectively, after two periods. Panel (e) illustrates a case when income distribution does not converge in the long run, but will cycle between Point A and Point B.

The more interesting income distribution dynamics are depicted in panels (b) and (d), in which different entry equilibria form for different initial income distributions, and there are two long run steady state income distributions. In panel (b), after one period, economies with an income distribution in the dark-shaded area reach the long-run distribution at Point C, while economies with a distribution in the light-shaded area reach the long-run distribution at Point A. Similarly, in panel (d), economies in different area of the income distribution space reach their respective long-run distributions, Points B and C, after one period. In both panels, starting from a more equal distribution (in the light-shaded area), an economy will reach a more equal steady state distribution, whereas starting from a more unequal distribution (in the dark-shaded area), an economy will reach a more unequal steady state distribution. Therefore, an initial equality or inequality is sustained, and it is sustained by an associated public education policy.

The scenario depicted in panel (d) of Fig. 1 is of particular interest because it is consistent with the stylized facts documented in Section 2. When income distributions are more unequal, the rich are the only lobby, and the education policy is more biased toward the rich — more spending on tertiary education that disproportionately benefits the rich. When income distributions are more equal, both the rich and the middle class lobby, and there is relatively less education spending on the rich and relatively more on the middle class — relatively less spending on tertiary education but more on secondary education that tends to benefit the middle class. Different initial income distributions are sustained through different public education policies.

4.4. Discussions

The above analysis is based on a very simplified model. I discuss the implications of modifying several assumptions of the model.

First, I have assumed that different individuals have different tastes for education. Alternatively, one may model differences among individuals by introducing a liquidity constraint: $F+C_i(p)<(1-\tau(p))w_i$. In other words, total spending on lobbying cannot exceed the net-of-tax income. While this constrained optimization problem may not be solved analytically, simulations show that the poor are less likely to enter lobbying and persistent differences in income distribution and education policy are more likely to emerge in the long run. In addition, with a binding liquidity constraint, a laissez-faire economy under-invests in education relative to the first best.

Second, I have assumed that socio-economic groups are of the same size because each family has one child. A long line of literature following Becker and Lewis (1973) and Becker and Tomes (1976) however has argued that high income is associated with low fertility and that parents substitute quality of children for quantity of children, where quality is taken to be human capital investment. Therefore, the rich may have fewer children but care more about their education, while the poor may have more children but care less about their education. We can easily incorporate this extension into the present model by interpreting individual i's valuation of his child's education δ_i as his total valuation for his children's education. The dynamics of education policy and income distribution can therefore be similarly analyzed as before.

Third, I have assumed that all socio-economic groups face the same entry cost. If, however, the rich have better network connections with government officials, or they have organized lobbies in pursuing other interests, then the cost of organizing an education lobby is incremental to them. Therefore, different socio-economic groups may face different entry cost in lobbying for *education*. In particular, the poor may face a much higher cost than others. Consequently, ceteris paribus, the poor are less likely to enter lobbying than the model has predicted.

Another extension of the present analysis is to relax the assumption that different public education policies have the same impact on economic growth. It is plausible that different allocations of public education spending have different ramifications for economic growth and be affected by different growth strategies. It would be instructive to explore whether there is an optimal human capital formation policy to achieve both economic growth and equality.

5. Conclusion

This paper provides a political economy explanation of the persistent differences in income distribution across countries over time. In contrast to the previous works that focus on total public education spending, the mechanism here is the allocation of public education spending on different schooling levels. Empirical regularities suggest that public spending on different schooling levels has distinct relationships with income distribution and is in general targeted at different socio-economic groups. Socio-economic groups may form lobbies to influence education policy making. The formation of lobbies is endogenous. Different lobbies may form in economies of different income distributions, but the rich are invariably more likely to lobby. When different lobbies form, public education policies will be different; over time, income distributions will evolve along different paths. In the long run, all economies may reach one unique steady state income distribution, or economies with different initial income distributions may reach different steady state distributions. In the latter case, we observe persistent equality or persistent inequality for different economies.

The endogeneity of lobby formation is essential to generate persistent differences in income distribution over time and multiple steady states in this model. In contrast, in a traditional median-voter model, policies always reflect the preferences of the median-voter, usually the median-income voter. With a right-skewed income distribution, it implies that all economies will converge to a relatively equal long run distribution. Similarly, a model with an exogenously given set of lobbies will also imply a unique dynamic path and a unique long-run income distribution. In both cases, changes in the economic and political parameters have limited roles in affecting the long-run distribution, whereas in the present model, changes in the economic or political parameters may have a profound impact on the long-run income distribution.

Appendix. Proofs of existence of equilibrium at entry stage

Appendix 1. Net-of-contribution utility

In this section, I directly derive j's equilibrium payoff let of contribution but gross of the fixed cost, $w_j=v_j-C_j$, and j's net payoff u_j is simply w_j-F.[24] In the derivation below, the superscript refers to the set of lobbies.

I use the results of Theorem 2 in Bernheim and Whinston (1986) to characterize the lobbies' net utilities. Theorem 2 states that in all truthful Nash equilibria, the government selects $p\in P^*$, and the lobbies receive payoffs in the set $E(P^*)$. $P^*=\text{argmax } kV+\sum_{i\in\mathcal{L}} v_i$ is

[24] To simplify notation, the superscript ^{o} that denotes equilibrium outcome is dropped.

the set of efficient policies that maximize the joint payoff of the lobbies and the government, and

$$E(p)\equiv\{w\in R^l|w\in\Pi(p)\text{ and there does not exist }\\ w'\in\Pi(p),\text{ with }w'\geq w\},$$

$$\Pi(p)\equiv\{w\in R^l|\text{for all }L\subseteq\mathcal{L},W_L\leq[kV(p)+V_{\mathcal{L}}(p)]\\ -[kV(p\overline{L})+V\overline{L}(p\overline{L})]\},$$

where l is the number of lobbies in $\mathcal{L}$, w is an l-dimensional vector of gross-of-fixed-cost payoffs to the lobbies in $\mathcal{L}$, $W_L=\sum_{i\in L} w_i$ $\overline{\mathrm{L}}$ is the complement of L in $\mathcal{L}$, $p\in P^*$, $p^{\overline{\mathrm{L}}}\in P^{\overline{\mathrm{L}}}$ with $P^{\overline{\mathrm{L}}}=\operatorname{argmax} kV+\sum_{i\in\overline{\mathrm{L}}} v_i$, $V_{\mathcal{L}}(p)=\sum_{i\in\mathcal{L}} v_i$ and $\mathrm{V}_{\overline{\mathrm{L}}}=\sum_{i\in\overline{\mathrm{L}}} v_i$. p and $\mathrm{p}^{\overline{\mathrm{L}}}$ can be found in Table 5 for different set of lobbies.

Case 1. j is the single lobby

Let $\mathcal{L}=\{j\}$.

$$w_j^{\mathcal{L}}\leq\left[kV\left(p^{\mathcal{L}}\right)+v_j\left(p^{\mathcal{L}}\right)\right]-kV\left(p^{\emptyset}\right).$$

Equality holds since j will maximize his net utility.[25]

Case 2. j is one of the two lobbies

Let $\mathcal{L}=\{i, j\}$, using the definition of $\Pi(p)$, we have

$$\begin{aligned}
&w_i^{\mathcal{L}}\leq\left[kV\left(p^{\mathcal{L}}\right)+v_i\left(p^{\mathcal{L}}\right)+v_j\left(p^{\mathcal{L}}\right)\right]-\left[kV\left(p^{\{j\}}\right)+v_j\left(p^{\{j\}}\right)\right]\\
&w_j^{\mathcal{L}}\leq\left[kV\left(p^{\mathcal{L}}\right)+v_i\left(p^{\mathcal{L}}\right)+v_j\left(p^{\mathcal{L}}\right)\right]-\left[kV\left(p^{\{i\}}\right)+v_i\left(p^{\{i\}}\right)\right]\\
&w_i^{\mathcal{L}}+w_j^{L}\leq\left[kV\left(p^{L}\right)+v_i\left(p^{\mathcal{L}}\right)+v_j\left(p^{\mathcal{L}}\right)\right]-kV\left(p^{\emptyset}\right)
\end{aligned}$$

The first two inequalities give the upper bounds on the values of $w_i^{\mathcal{L}}$ and $w_j^{\mathcal{L}}$. They give the exact values of $w_j^{\mathcal{L}}$ and $w_i^{\mathcal{L}}$ if the third inequality is redundant. The third inequality can be shown to be redundant when $\mathcal{L}=\{1, 2\}$; or when $\mathcal{L}=\{1, 3\}$, and the difference between δ_2 and δ_3 is not excessively large; or when $\mathcal{L}=\{2, 3\}$, and the difference between δ_1 and δ_2 is not excessively large.

Case 3. j is one of the three lobbies

Let $\mathcal{L}=\{1, 2, 3\}$ and $s_1\geq s_2\geq s_3$, we have

$$\begin{aligned}
&w_1^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{2,3\}}\right)+v_2\left(p^{\{2,3\}}\right)+v_3\left(p^{\{2,3\}}\right)\right]\\
&w_1^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{1,3\}}\right)+v_1\left(p^{\{1,3\}}\right)+v_3\left(p^{\{1,3\}}\right)\right]\\
&w_1^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{1,2\}}\right)+v_1\left(p^{\{1,2\}}\right)+v_2\left(p^{\{1,2\}}\right)\right]\\
&w_1^{\mathcal{L}}+w_2^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{3\}}\right)+v_3\left(p^{\{3\}}\right)\right]\\
&w_1^{\mathcal{L}}+w_3^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{2\}}\right)+v_2\left(p^{\{2\}}\right)\right]\\
&w_2^{\mathcal{L}}+w_3^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-\left[kV\left(p^{\{1\}}\right)+v_1\left(p^{\{1\}}\right)\right]\\
&w_1^{\mathcal{L}}+w_2^{\mathcal{L}}+w_3^{\mathcal{L}}\leq(k+1)V\left(p^{\mathcal{L}}\right)-kV\left(p^{\emptyset}\right)
\end{aligned}$$

The first three inequalities give the upper bounds on the values of $w_1^{\mathcal{L}}$, $w_2^{\mathcal{L}}$, and $w_3^{\mathcal{L}}$. Since $p^{\mathcal{L}}=p^{\emptyset}$, the last inequality is redundant. The fourth inequality is redundant if $k\geq1$; otherwise, it is redundant if the δ's are lot too close to each other. The fifth inequality is redundant if the difference between δ_2 and δ_3 is not excessively large. The sixth inequality is redundant if differences between δ's are not excessively large.

In what follows, I use these upper bounds as values of w's in the two-lobby and three-lobby cases to derive the conditions for each of the entry equilibria. This practice is harmless in proving the four lemmas below for two reasons. First, since entry equilibria are characterized by inequalities, using these upper bounds frequently imposes more stringent conditions for most entry equilibria. Second, as discussed above, with loose restrictions on parameter values (δ_i's ald k), these upper bounds provide the exact net utility values.

Appendix 2. Conditions for possible entry equilibria

There are eight possible pure strategy equilibria at the entry stage. The conditions for each of the equilibria are as follows.

Conditions for $\mathcal{L}=\emptyset$

$\forall i\in\mathcal{I}$, the net utility when entering is smaller than that when not given that the other two individuals do not enter.

$$\left\{s_1^2\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-s_1\left(2\delta_1^2+F\right)+\left(\delta_1^2-k\cdot F\right)\right\}/\left(k+s_1\right)<0 \tag{1a}$$

$$\left\{s_2^2\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-s_2\left(2\delta_2^2+F\right)+\left(\delta_2^2-k\cdot F\right)\right\}/\left(k+s_2\right)<0 \tag{1b}$$

$$\left\{s_3^2\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-s_3\left(2\delta_3^2+F\right)+\left(\delta_3^2-k\cdot F\right)\right\}/\left(k+s_3\right)<0 \tag{1c}$$

Conditions for $\mathcal{L}=\{1\}$

1 receives a higher net utility when entering than not, given that 2 and 3 do not enter; given that 1 enters and 3 (2) does not, 2(3) receives a lower utility when entering than not.

$$\mathrm{LHS}(1a)>0 \tag{2a}$$

$$\begin{aligned}\{&s_2^2\left(\delta_1^2(k+1)^2+\left(\delta_2^2+\delta_3^2\right)k^2\right)-s_2(k+s_1)\left(2k\delta_2^2+F(k+s_1)\right)\\ &+(k+s_1)^2\left(\delta_2^2-F(k+s_1)\right)\}/\left((k+s_1+s_2)(k+s_1)^2\right)<0\end{aligned} \tag{2b}$$

[25] The expression for $w_j^{\mathcal{L}}$ can also be derived directly from the definition of SPNE. In equilibrium, Eq. (5) is binding, and $W(p^{\mathcal{L}}, C_j^o(p^{\mathcal{L}}))=W(p^{\emptyset},0)$. Since $W(p^{\mathcal{L}}, C_j^o(p^{\mathcal{L}}))=kV(p^{\mathcal{L}})+C_j^o(p^{\mathcal{L}})-kF$ and $W(p^{\emptyset},0)=kV(p^{\emptyset})-kF$, we have $C_j^o(p^{\mathcal{L}})=kV(p^{\emptyset})-kV(p^{\mathcal{L}})$, and thus the above equality.

$$\{s_3^2\left(\delta_1^2(k+1)^2+\left(\delta_2^2+\delta_3^2\right)k^2\right)-s_3(k+s_1)\left(2k\delta_3^2+F(k+s_1)\right)+(k+s_1)^2\left(\delta_3^2-F(k+s_1)\right)\}/\left((k+s_1+s_3)(k+s_1)^2\right)<0 \tag{2c}$$

Conditions for $\mathcal{L}$ ={2}

2 receives a higher net utility when entering than not, given that 1 and 3 do not enter; given that 2 enters and 3(1) does not, 1(3) receives a lower utility when entering than not.

$$\mathrm{LHS}(1b)>0 \tag{3a}$$

$$\{s_1^2\left(\delta_2^2(k+1)^2+\left(\delta_1^2+\delta_3^2\right)k^2\right)-s_1(k+s_2)\left(2k\delta_1^2+F(k+s_2)\right)+(k+s_2)^2\left(\delta_1^2-F(k+s_2)\right)\}/\left((k+s_1+s_2)(k+s_2)^2\right)<0 \tag{3b}$$

$$\{s_3^2\left(\delta_2^2(k+1)^2+\left(\delta_1^2+\delta_3^2\right)k^2\right)-s_3(k+s_2)\left(2k\delta_3^2+F(k+s_2)\right)+(k+s_2)^2\left(\delta_3^2-F(k+s_2)\right)\}/\left((k+s_2+s_3)(k+s_2)^2\right)<0 \tag{3c}$$

Conditions for $\mathcal{L}$={3}

3 receives a higher net utility when entering than not, given that 1 and 2 do not enter; Given that 3 enters and 2(1) does not, 1(2) receives a lower utility when entering than not.

$$\mathrm{LHS}(1c)>0 \tag{4a}$$

$$\{s_1^2\left(\delta_3^2(k+1)^2+\left(\delta_1^2+\delta_2^2\right)k^2\right)-s_1(k+s_3)\left(2k\delta_1^2+F(k+s_3)\right)+(k+s_3)^2\left(\delta_1^2-F(k+s_3)\right)\}/\left((k+s_1+s_3)(k+s_3)^2\right)<0 \tag{4b}$$

$$\{s_2^2\left(\delta_3^2(k+1)^2+\left(\delta_1^2+\delta_2^2\right)k^2\right)-s_2(k+s_3)\left(2k\delta_2^2+F(k+s_3)\right)+(k+s_3)^2\left(\delta_2^2-F(k+s_3)\right)\}/\left((k+s_2+s_3)(k+s_3)^2\right)<0 \tag{4c}$$

Conditions for $\mathcal{L}$={1, 2}

1(2) receives a higher net utility when entering than not, given that 2(1) enters and 3 does not; 3 receives a lower net utility when entering than not, given that both 1 and 2 enter.

$$\mathrm{LSH}(3b)>0 \tag{5a}$$

$$\mathrm{LSH}(2b)>0 \tag{5b}$$

$$\{s_3^2\left((k+1)\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-F\right)-2s_3(k+1)\left(\delta_3^2-F\right)+(k+1)\left(\delta_3^2-F(k+1)\right)\}/\left((k+s_1+s_2)^2\right)<0 \tag{5c}$$

Conditions for $\mathcal{L}$={1, 3}

1(3) receives a higher net utility when entering than not, given that 3(1) enters and 2 does not; 2 receives a lower net utility when entering than not, given that both 1 and 3 enter.

$$\mathrm{LSH}(4b)>0 \tag{6a}$$

$$\mathrm{LSH}(2c)>0 \tag{6b}$$

$$\{s_2^2\left((k+1)\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-F\right)-2s_2(k+1)\left(\delta_2^2-F\right)+(k+1)\left(\delta_2^2-F(k+1)\right)\}/\left((k+s_1+s_3)^2\right)<0 \tag{6c}$$

Conditions for $\mathcal{L}$={2, 3}

2(3) receives a higher net utility when entering than not, given that 3(2) enters and 1 does not; 1 receives a lower net utility when entering than not, given that both 2 and 3 enter.

$$\mathrm{LSH}(4c)>0 \tag{7a}$$

$$\mathrm{LSH}(3c)>0 \tag{7b}$$

$$\{s_1^2\left((k+1)\left(\delta_1^2+\delta_2^2+\delta_3^2\right)-F\right)-2s_1(k+1)\left(\delta_1^2-F\right)+(k+1)\left(\delta_1^2-F(k+1)\right)\}/\left((k+s_2+s_3)^2\right)<0 \tag{7c}$$

Conditions for $\mathcal{L}$={1, 2, 3}

$\forall i\in\mathcal{I}$, the net utility when entering is higher than that when lot, given that others enter.

$$\mathrm{LHS}(7c)>0 \tag{8a}$$

$$\mathrm{LHS}(6c)>0 \tag{8b}$$

$$\mathrm{LHS}(5c)>0 \tag{8c}$$

The denominator of the left-hand side of each of the inequalities (1a)–(8c) is positive. In the proofs below, I consider only the numerator, which is a quadratic function of the corresponding income share.

Appendix 3. Proofs

Definition A-1. *The discriminant of a quadratic equation $f(x)=ax^2+bx+c=0$ is defined as $\Delta=b^2-4ac$.*

A quadratic equation has no real root if $\Delta<0$; it has two identical real roots if $\Delta=0$; it has two different real roots if $\Delta>0$. When $a>0$, $\Delta\le 0$ is equivalent to $f(x)\ge 0$ on its domain. When $a>0$ and $f(x)=0$ has two different real roots x^{low} and x^{high}, $f(x)>0$ for $x<x^{\mathrm{low}}$ and $x>x^{\mathrm{high}}$.

I prove the results in Section 4.2 by proving the following lemmas; they correspond to scenarios with increasing entry costs. The restrictions on $D_{12}=\delta_1^2/\delta_2^2$ and $D_{23}=\delta_2^2/\delta_2^2$ insure the existence and uniqueness of a pure strategy equilibrium for any given initial income distribution for the values of entry cost specified in each lemma. These restrictions can be relaxed significantly if we do not consider extreme income distributions. In the proofs, whenever an equation number is referred to, it means the numerator of the left-hand side of the corresponding inequality in Appendix-2.

Lemma A1. *Let* $F_0=\frac{\delta_3^2(k+1)\left(\delta_1^2+\delta_2^2\right)}{(k+1)^2\left(\delta_1^2+\delta_2^2\right)+k^2\delta_3^2}$. *If* $F<F_0$, *then* $\forall s\in S^3$, $\mathcal{L}=\{1, 2, 3\}$ *is an equilibrium.*

Proof. $F_0=\{F|\varDelta_{(8c)}=0\}$. When $F<F_0$, $\varDelta_{(8a)}<\varDelta_{(8b)}<\varDelta_{(8c)}<0$, hence Eq. (8a)>0, Eq. (8b)>0, and Eq. (8c)>0. □

Lemma A2. *If* $D_{23}\geq\frac{2k+1}{k+1}$ *and* $D_{12}\geq max\left\{1+\frac{1}{4k(2k+1)},\frac{8(k+1)^2+D_{23}\left(8k^2+1\right)}{16(k+1)^2+D_{23}\left(8k^2-8k\right)}\right\}$, *then for* $F_0\leq F\leq\frac{\delta_2^2}{k+1}$ $\forall s\in S^3$,

(i) A unique equilibrium exists;

(ii) the equilibrium is either $\mathcal{L}=\{1, 2, 3\}$ *or* $\mathcal{L}=\{1, 2\}$. *Moreover, for* $F<F_0$, $\mathcal{L}=\{1, 2, 3\}$ *is the unique equilibrium for all* $s\in S^3$.

Proof. I prove the lemma by showing that, regardless of his income, individual 1's dominant strategy is to enter; given that individual 1 enters irrespective of others' actions, individual 2's iterated dominant strategy is also to enter regardless of his income; individual 3's strategy, in contrast, depends upon his income, as well as specific values of the model parameters.

(1) Individual 1's dominant strategy is to enter, regardless of his income.

First, $D_{23}\geq\frac{2k+1}{k+1}$ is sufficient for $F_0<\frac{\delta_2^2}{k+1}<\frac{\delta_2^2}{2k+1}<F_1$ where $F_1=\frac{\delta_2^2(k+1)\left(\delta_1^2+\delta_3^2\right)}{(k+1)^2\left(\delta_1^2+\delta_3^2\right)+k^2\delta_2^2}$. Since $F_1=\{F|\varDelta_{(8b)}=0\}$, for $F\leq\frac{\delta_2^2}{2k+1}$, $\varDelta_{(8a)}<\varDelta_{(8b)}<0$, hence Eq. (8a)>0. Given that 2 and 3 enter, 1 will enter.

Second, $\varDelta_{(5a)}<0$ for $F\leq\frac{\delta_2^2}{2k+1}$, hence Eq. (5a)>0. To see this,

$$
\begin{aligned}
\varDelta_{(5a)}=&\left\{F^2(k+s_2)^2+4F(k+s_2)\left(\delta_1^2k(k+1)+\delta_2^2(k+1)^2+\delta_3^2k^2\right)\right.\\
&\left.-4\delta_1^2\left(\delta_2^2(k+1)^2+\delta_3^2k^2\right)\right\}(k+s_2)^2\\
\leq&\left\{F^2\left(k+\frac{1}{2}\right)^2+4F\left(k+\frac{1}{2}\right)\left(\delta_1^2k(k+1)+\delta_2^2(k+1)^2+\delta_3^2k^2\right)\right.\\
&\left.-4\delta_1^2\left(\delta_2^2(k+1)^2+\delta_3^2k^2\right)\right\}(k+s_2)^2\\
\leq&\left\{\delta_2^4+8\delta_2^2(\delta_1^2k(k+1)+\delta_2^2(k+1)^2+\delta_3^2k^2)\right.\\
&\left.-16\delta_1^2\left(\delta_2^2(k+1)^2+\delta_3^2k^2\right)\right\}(k+s_2)^2\\
<&0.
\end{aligned}
$$

Dividing both sides by δ_2^2 and δ_3^2, it is equivalent to

$$
D_{12}>\frac{D_{23}\left(1+8(k+1)^2+8k^2\right)}{D_{23}\left(8(k+1)^2+8(k+1)\right)+16k^2}=\text{RHS},
$$

which holds since $D_{12}\geq 1>RHS$. Given that 2 enters and 3 does not, 1 will enter.

Third, $D_{12}>\frac{8(k+1)^2+D_{23}\left(8k^2+1\right)}{16(k+1)^2D_{23}\left(8k^2-8k\right)}$ is sufficient for $\varDelta_{(6a)}<0$, hence (6*a*)>0. To see this,

$$
\begin{aligned}
\varDelta_{(6a)}=&\left\{F^2(k+s_3)^2+4F(k+s_3)\left(\delta_1^2k(k+1)+\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right.\\
&\left.-4\delta_1^2\left(\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right\}(k+s_3)^2\\
\leq&\left\{F^2\left(k+\frac{1}{2}\right)^2+4F\left(k+\frac{1}{2}\right)\left(\delta_1^2k(k+1)+\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right.\\
&\left.-4\delta_1^2\left(\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right\}(k+s_3)^2\\
\leq&\left\{\delta_2^4+8\delta_2^2\left(\delta_1^2k(k+1)+\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right.\\
&\left.-16\delta_1^2\left(\delta_2^2k^2+\delta_3^2(k+1)^2\right)\right\}(k+s_3)^2\\
<&0
\end{aligned}
$$

is equivalent to

$$
D_{12}>\frac{8(k+1)^2+D_{23}(8k^2+1)}{16(k+1)^2+D_{23}(8k^2-8k)}.
$$

Given that 3 enters and 2 does not, 1 will enter.

Last, $D_{12}\geq 1+\frac{1}{4k(2k+1)}$ is sufficient for $\varDelta_{(2a)}<0$, hence Eq. (2a)>0. To see this,

$$
\begin{aligned}
\varDelta_{(2a)}=&F^2+4F\left((k+1)\delta_1^2+k\delta_2^2+k\delta_3^2\right)-4\delta_1^2\left(\delta_2^2+\delta_3^2\right)\\
\leq&\frac{\delta_2^4}{(2k+1)^2}+\frac{4\delta_2^2}{2k+1}\left((k+1)\delta_1^2+k\delta_2^2+k\delta_3^2\right)-4\delta_1^2\left(\delta_2^2+\delta_3^2\right)\\
<&0
\end{aligned}
$$

is equivalent to

$$
D_{12}>\frac{4k(2k+1)+D_{23}(4k(2k+1)+1)}{4(2k+1)^2+D_{23}(4k(2k+1))},
$$

which holds if $D_{12}>1+\frac{1}{4k(2k+1)}$. Given that neither 2 nor 3 enters, 1 will enter.

(2) Given that individual 1 enters irrespective of others' actions, individual 2's iterated dominant strategy is to enter regardless of his income.

First, For $F\leq\frac{\delta_2^2}{2k+1}$, $\varDelta_{(8b)}>0$, hence Eq. (8b)>0. Given that 1 and 3 enter, 2 will enter.

Second, $D_{12} \geq 1 + \frac{1}{4k(2k+1)}$ is sufficient for $\Delta_{(5b)} < 0$, hence Eq. (5b) > 0. To see this,

$$\begin{aligned}\Delta_{(5b)} = &\Big\{F^2(k+s_1)^2 + 4F(k+s_1)\left(\delta_1^2(k+1)^2 + \delta_2^2 k(k+1) + \delta_3^2 k^2\right)\\ &-4\delta_2^2\left(\delta_1^2(k+1)^2 + \delta_3^2 k^2\right)\Big\}(k+s_1)^2\\ \leq &\Big\{F^2(k+1)^2 + 4F(k+1)\left(\delta_1^2(k+1)^2 + \delta_2^2 k(k+1) + \delta_3^2 k^2\right)\\ &-4\delta_2^2\left(\delta_1^2(k+1)^2 + \delta_3^2 k^2\right)\Big\}(k+s_1)^2\\ \leq &\left\{\frac{\delta_2^4(k+1)^2}{(2k+1)^2} + \frac{4\delta_2^2(k+1)}{2k+1}\left(\delta_1^2(k+1)^2 + \delta_2^2 k(k+1) + \delta_3^2 k^2\right)\right.\\ &\left.-4\delta_2^2\left(\delta_1^2(k+1)^2 + \delta_3^2 k^2\right)\right\}(k+s_1)^2\\ < &\,0\end{aligned}$$

is equivalent to

$$D_{12} > 1 + \frac{1}{4k(2k+1)} - \frac{k^2}{D_{23}(k+1)^2},$$

which holds if $D_{12} > 1 + \frac{1}{4k(2k+1)}$. Given that 1 enters and 3 does not, 2 will enter.

(3) Individual 3's action depends on his income, as well as specific values of the model parameters. (8c) = 0 has two real roots: s_3^{low} and s_3^{high}.
- If $0 < s_3^{\text{low}} < s_3^{\text{high}} < \frac{1}{3}$, then Eq. (8c) > 0 for $s_3 \in (0, s_3^{\text{low}}) \cup (s_3^{\text{high}}, \frac{1}{3})$; given that 1 and 2 enter, 3 will enter. Otherwise, 3 will not enter.
- If $s_3^{\text{low}} < 0$ and $s_3^{\text{high}} > \frac{1}{3}$, then Eq. (8c) < 0 for all feasible values of s_3; given that 1 and 2 enter, 3 will not enter.
- If $s_3^{\text{low}} < 0$ and $s_3^{\text{high}} < \frac{1}{3}$, then (8c) > 0 for all $s_3 \in (s_3^{\text{high}}, \frac{1}{3})$; given that 1 and 2 enter, 3 will enter. Otherwise, 3 will not enter.
- If $s_3^{\text{low}} > 0$ and $s_3^{\text{high}} > \frac{1}{3}$, then Eq. (8c) > 0 for all $s_3 \in (0, s_3^{\text{low}})$; given that 1 and 2 enter, 3 will enter. Otherwise, 3 will not enter.

(4) For $F < F_0$, steps (1) and (2) hold. Moreover, individual 3 also has a dominant strategy – to enter – regardless of his income. Therefore, the unique equilibrium is $\mathcal{L} = \{1, 2, 3\}$. □

Lemma A3. *If max* $\left\{1.5, \frac{2k+1}{k+1}\right\} \leq D_{23} \leq \widehat{D_{23}}$, *where* $\widehat{D_{23}}$ *is a decreasing function of k and* $\lim_{k\to\infty} \widehat{D_{23}} = 3.732$, *and* $D_{12} \geq \max\left\{\frac{(k+1)(D_{23}+1)}{D_{23}+k+1}, \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}\right\}$, *then for* $\frac{\delta_2^2}{2k+1} < F \leq \frac{\delta_2^2}{k+1}$, $\forall s \in S^3$,

(i) An equilibrium exists;
(ii) the equilibrium is either $\mathcal{L} = \{1, 2, 3\}$, $\mathcal{L} = \{1, 2\}$, $\mathcal{L} = \{1, 3\}$, *or* $\mathcal{L} = \{1\}$;
(iii) whenever $\mathcal{L} = \{1, 3\}$ *is an equilibrium, it overlaps with the equilibrium* $\mathcal{L} = \{1, 2\}$.

Proof. I prove the lemma by showing that individual 1 has a dominant strategy – to enter – regardless of his income, while the actions of individuals 2 and 3 depend upon their income, as well as specific values of the model parameters.

(1) Individual 1's dominant strategy is to enter, regardless of his income.

First, $D_{12} \geq \frac{(k+1)(D_{23}+1)}{k+D_{23}+1}$ is sufficient for $F_1 < \frac{\delta_2^2}{k+1} < F_2$, where $F_2 = \frac{\delta_1^2(k+1)\left(\delta_2^2+\delta_3^2\right)}{(k+1)^2\left(\delta_2^2+\delta_3^2\right)+k^2\delta_1^2}$. Since $F_2 = \{F | \Delta_{(8a)} = 0\}$, for $\frac{\delta_2^2}{2k+1} < F \leq \frac{\delta_2^2}{k+1}$ $\Delta_{(8a)} < 0$, hence Eq. (8a) > 0. Given that 2 and 3 enter, 1 will enter.

Second, $D_{12} > \frac{(k+1)(D_{23}+1)}{k+D_{23}+1}$ is sufficient for $\Delta_{(5a)} < 0$, hence Eq. (5a) > 0. To see this,

$$\begin{aligned}\Delta_{(5a)} \leq &\left\{\frac{\delta_2^4(2k+1)^2}{4(k+1)^2} + \frac{2\delta_2^2(2k+1)}{k+1}\left(\delta_2^2 k(k+1) + \delta_2^2(k+1)^2 + \delta_3^2 k^2\right)\right.\\ &\left.-4\delta_1^2\left(\delta_2^2(k+1)^2 + \delta_3^2 k^2\right)\right\}(k+s_2)^2\\ < &\,0\end{aligned}$$

is equivalent to

$$\begin{aligned}D_{12} > &\frac{8D_{23}(k+1)^3(2k+1) + D_{23}(2k+1)^2 + 8(2k+1)(k+1)k^2}{8D_{23}(k+1)^2(3k+2) + 16k^2(k+1)^2}\\ = &\,\text{RHS}.\end{aligned}$$

It holds since $D_{12} > \frac{(k+1)(D_{23}+1)}{k+D_{23}+1} > \text{RHS}$. Given that 2 enters and 3 does not, 1 will enter.

Third, $D_{12} > \max\left\{1, \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}\right\}$ is sufficient for $\Delta_{(2a)} < 0$, hence Eq. (2a) > 0.To see this,

$$\begin{aligned}\Delta_{(2a)} \leq &\frac{\delta_2^4}{(k+1)^2} + \frac{4\delta_2^2}{k+1}\left((k+1)\delta_1^2 + k\delta_2^2 + k\delta_3^2\right)\\ &-4\delta_1^2\left(\delta_2^2 + \delta_3^2\right)\\ < &\,0\end{aligned}$$

is equivalent to

$$D_{12} > \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2} = \text{RHS}.$$

The RHS may or may not be greater than 1. Thus, $D_{12} > \max\left\{1, \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}\right\}$. Given that neither 2 nor 3 enters, 1 will enter.

Last, as long as D_{23} is not too large, $D_{12} > \max\left\{1, \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}\right\}$ is sufficient for $\Delta_{(6a)} < 0$, hence Eq. (6a) > 0. To see this,

$$\begin{aligned}\Delta_{(6a)} \leq &\left\{\frac{\delta_2^4(3k+1)^2}{9(k+1)^2} + \frac{4\delta_2^2(3k+1)}{3(k+1)}\left(\delta_1^2 k(k+1) + \delta_2^2 k^2 + \delta_3^2(k+1)^2\right)\right.\\ &\left.-16\delta_1^2\left(\delta_2^2 k^2 + \delta_3^2(k+1)^2\right)\right\}(k+s_3)^2\\ < &\,0\end{aligned}$$

is equivalent to

$$D_{12} > \frac{12(k+1)^3(3k+1) + D_{23}\left(12k^2(3k+1)(k+1) + (3k+1)^2\right)}{12(k+1)^2\left(3(k+1)^2 - kD_{23}\right)}$$
$$= \text{RHS}.$$

It can be shown that $\frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}$>RHS as long as $D_{23} < \widehat{D_{23}(k)}$, where $\widehat{D_{23}(k)}$ is a decreasing function of k, and $\lim_{k\to\infty}\widehat{D_{23}(k)} = 3.732$. Given that 3 enters and 2 does not, 1 will enter.

(2) The actions of individual 2 and 3 depend on their income, as well as specific values of the model parameters.

First, Eq. (8c)=0 has two roots of opposite signs: s_3^+ and s_3^-.

– If $s_3^+ \geq \frac{1}{3}$, then (8c)<0 for all $s_3 \in (0, \frac{1}{3})$.$\mathcal{L} = \{1, 2, 3\}$ is lot al equilibrium.
– If $s_3^+ < \frac{1}{3}$, then Eq. (8c)>0 for $s_3 \in (s_3^+, \frac{1}{3})$; 3 will enter given that 1 and 2 do.

• For $F < F_1$, $\Delta_{(8b)} < 0$ and Eq. (8b)>0.
• For $F > F_1$, $\Delta_{(8b)} > 0$ and Eq. (8b)=0 has two positive roots: s_2^{high} and s_2^{low}.

Since Eqs. (8b) and (8c) belongs to the same equation family

$$G = Ax^2 - 2(k+1)(\delta^2 - F)x + (k+1) \times (\delta^2 - F(k+1)) = 0,$$

where $A = (k+1)(\delta_1^2 + \delta_2^2 + \delta_3^2) - F$; and

$$\frac{\partial x^{\text{high}}}{\partial \delta^2} = -\frac{\partial G/\partial \delta^2}{\partial G/\partial x^{\text{high}}} = -\frac{(k+1)(1 - 2x^{\text{high}})}{\sqrt{\Delta_G}} < 0$$

if $x^{\text{high}} < \frac{1}{2}$.
Since $s_3^+ < \frac{1}{3}$, for δ_2^2 sufficiently close to δ_3^2, $s_2^{\text{high}} < s_3^+$. Moreover, $\frac{\partial s_2^{\text{high}}(8b)}{\partial \delta_2^2} < 0$. Therefore, for all $\delta_2^2 > \delta_3^2$, we have $s_2^{\text{high}} < s_3^+ < \frac{1}{3}$. Since $s_2 \geq s_3 > s_3^+ > s_2^{\text{high}}$, Eq. (8b)>0.

Thus for $s_3 \geq \in (s_3^+, \frac{1}{3})$, $\mathcal{L} = \{1, 2, 3\}$ is an equilibrium.

Second, consider the subspace of S^3 with $s_3 < s_3^+$, such that Eq. (8c)<0, or equivalently, Eq. (5c)<0. Moreover, we know Eq. (5a)>0 and Eq. (2a)>0.

– Since $\Delta_{(5b)}$ increases with s_1, $\exists s_1^*$, such that when $s_1 < s_1^*$, $\Delta_{(5b)} < 0$; and when $s_1 \geq s_1^*$, $\Delta_{(5b)} > 0$. If $s_1^* > 1$, then $\Delta_{(5b)} < 0$ for all s in this subspace of S^3, hence Eq. (5b)>0, and {1, 2} is an equilibrium. If $s_1^* < 1$, then for $\frac{1}{3} < s_1 < s_1^*$, $\Delta_{(5b)} < 0$, hence Eq. (5b)>0, and {1, 2} is an equilibrium; for $s_1^* < s_1 < 1$, Eq. (5b)=0 partitions the relevant space into two parts: Eq. (5b)<0 for $s_2 \in (s_2^{\text{low}}, s_2^{\text{high}})$, and Eq. (5b)>0 hence {1, 2} is an equilibrium, for otherwise.
– Consider the subspace $s_2 \in (s_2^{\text{low}}, s_2^{\text{high}})$, such that Eq. (5b)<0, or equivalently, Eq. (2b)<0. I show that Eq. (2c)<0 for this region.

• Since $\Delta_{(2c)} > \Delta_{(2b)}$ when $\delta_1^2(k+1)^2 > Fk(k+s_1)$, which holds trivially for $F < \frac{\delta_2^2}{k+1}$, $\Delta_{(2c)} > 0$. Thus, Eq. (2c)=0 has two real roots: s_3^{low} and s_3^{high}. Eq. (2c)<0 is equivalent to $s_3^{\text{low}} < s_3 < s_3^{\text{high}}$.
• $s_3 < s_3^{\text{high}}$ if $s_2 < s_2^{\text{high}}$. To see this, note that Eqs. (2b) and (2c) belong to the same equation family

$$G = Ax^2 - (k+s_1)\left(2k\delta^2 + F(k+s_1)\right)x + (k+s_1)^2\left(\delta^2 - F(k+s_1)\right) = 0$$

where $A = \delta_1^2(k+a)^2 + \delta_2^2 k^2 + \delta_3^2 k^2$; and

$$\frac{\partial x^{\text{high}}}{\partial \delta^2} = \frac{\partial G/\partial \delta^2}{\partial G/\partial x^{\text{high}}} = -\frac{(k+s_1)(k+s_1-2kx^{\text{high}})}{\sqrt{\Delta_G}} < 0$$

if $x^{\text{high}} < \frac{1}{2}$. Moreover, $\frac{\partial s_2^{\text{high}}(2b)}{\partial \delta_2^2} < 0$.
Therefore, if $s_3^{\text{high}} < \frac{1}{2}$, then $s_2^{\text{high}} < s_3^{\text{high}} < \frac{1}{2}$ for all $\delta_2^2 > \delta_3^2$. Since $s_2 < s_2^{\text{high}}$ and $s_3 \leq s_2$, $s_3 < s_3^{\text{high}}$. If, on the contrary, $s_3^{\text{high}} > \frac{1}{2}$, then $s_3 < s_3^{\text{high}}$ by definition.
• To show $s_3 > s_3^{\text{low}}$, it is sufficient if $s_3^{\text{low}} < 0$. Since $s_3^{\text{low}} + s_3^{\text{high}} > 0$, we need $s_3^{\text{low}} \cdot s_3^{\text{high}} < 0$; this in turn sufficient if $\delta_3^2 < F(k+s_1)$, or equivalently, $s_1 > \frac{\delta_3^2}{F} - k = s_1^\star$. Since the relevant subspace exhibits $s_1^* < s_1 < 1$, we want to show $s_1^\star < 1$ and $s_1^\star < s_1^*$. First,

$$s_1^\star = \frac{\delta_3^2}{F} - k < \frac{\delta_3^2}{\delta_2^2(2k+1)} - k < (2k+1)\frac{k+1}{2k+1} - k = 1$$

The first inequality holds because $F > \frac{\delta_2^2}{2k+1}$, and the second because $\frac{\delta_2^2}{\delta_3^2} > \frac{2k+1}{k+1}$. Second, $s_1^\star < s_1^*$ is equivalent to $\Delta_{(2b)}(s_1^\star) < 0$, which, after tedious manipulation, is equivalent to $D_{12} > \frac{1+4k^2+4kD_{23}}{4(k+1)^2\left(D_{23}^2 - D_{23}\right)}$. When $D_{12} > max\left\{1, \frac{k(D_{23}+1)}{k+1} + \frac{D_{23}}{4(k+1)^2}\right\}$, it holds as long as D_{23} is not too small. A sufficient condition is $D_{23} \geq \max\left\{1.5, \frac{2k+1}{k+1}\right\}$.

Thus $\mathcal{L} = \{1\}$ is and equivalent for this region.

(3) In sum, in equivalent, $\forall s \in S^3$, an equivalent exist, and it can be either $\mathcal{L} = \{1, 2, 3\}$, $\mathcal{L} = \{1, 2\}$, or $\mathcal{L} = \{1\}$. We cannot however rule out $\mathcal{L} = \{1, 3\}$ as an equilibrium, but we know that whenever it arises, it has to overlap with $\mathcal{L} = \{1, 2\}$, since it is mutually exclusive with $\mathcal{L} = \{1, 2, 3\}$ and $\mathcal{L} = \{1\}$. □

Lemma A4. *Let* $\widehat{D_{12}} = min\left\{3 + \frac{2}{k} - \frac{1}{D_{23}}, 8 + \frac{k-1}{k^2} - \frac{4}{D_{23}}\right\}$. *For* $F \geq \frac{\delta_2^2}{k}$ *if k is sufficiently large, and* $D_{23} > 1$, $\widehat{D_{12}} \geq D_{12} > 1$, *then* $\forall s \in S^3$,

(i) A unique equilibrium exists;
(ii) the equilibrium is either $\mathcal{L} = \{1, 2\}$, $\mathcal{L} = \{1\}$, *or* $L = \emptyset$.

Proof. I prove the lemma by showing that individual 3's dominant strategy is to not enter regardless of his income, while the actions of individual 1 and 2 depend on their income, as well as specific values of the model parameters.

(1) Individual 3's dominant strategy is to not enter regardless of his income, i.e., $\forall s_3 \in \left[0, \frac{1}{3}\right]$, Eqs. (1c)<0, (2c)<0, (3c)<0 and (8c)<0 when $F \geq \frac{\delta_2^2}{k}$. It is obvious that Eqs. (1c)=0, (2c)=0, (3c)=0, and (8c)=0 each have two real roots of opposite signs when $F \geq \frac{\delta_2^2}{k}$, and the positive roots increase with F; therefore, it is sufficient to show the positive roots are greater than $\frac{1}{3}$.

First, $D_{12} \leq 8 + \frac{3}{k} - \frac{4}{D_{23}} = \widehat{D_a}$ is sufficient for $s_3^+(1c) \geq \frac{1}{3}$.
Second, a sufficient condition for $s_3^+(2c) \geq \frac{1}{3}$ is that $s_3^+(s_1 = \frac{1}{3})\ (2c) \geq \frac{1}{3}$. This is because

$$\frac{\partial s_3^+(2c)}{\partial s_1} = -\frac{\partial(2c)/\partial s_1}{\partial(2c)/\partial s_3^+(2c)} = \frac{3F\left(k^2+s_1^2\right)+2Fk+2s_1\left(3kF+F-\delta_3^2\right)}{\sqrt{\Delta_{(2c)}}} > 0$$

when $F \geq \frac{\delta_2^2}{k}$. Therefore, $D_{12} \leq \frac{(3k+1)^2(3k+2)-3k^3}{3k(k+1)^2} - \frac{(2k+1)^2}{(k+1)^2 D_{23}} = \widehat{D_b}$ is sufficient for (2c)<0.
Third, similarly, a sufficient condition for $s_3^+(3c) \geq \frac{1}{3}$ is that $s_3^+(s_2 = 0)(3c) \geq \frac{1}{3}$, which is equivalent to $D_{12} \leq 8 + \frac{k-1}{k^2} - \frac{4}{D_{23}} = \widehat{D_c}$.[26]
Last, $D_{12} \leq \frac{(3k+2)^2}{k(k+1)} - 1 - \frac{4}{D_{23}} = \widehat{D_d}$ is sufficient for $s_3^+(8c) \geq \frac{1}{3}$.

(2) It is obvious that Eq. (1b)=0 has two real roots of opposite signs when $F \geq \frac{\delta_2^2}{k}$, and its positive root increases with F. Thus we can show Eq. (1b)<0 by showing $s_2^+\left(F = \frac{\delta_2^2}{k}\right)(1b) \geq \frac{1}{2}$. A sufficient condition is $D_{12} \leq 3 + \frac{2}{k} - \frac{1}{D_{23}} = \widehat{D}_e$. In other words, with more stringent constraint on D_{12}, we have that 2 will not enter given that 1 and 3 do not.

(3) There are only three possible equilibria: $\emptyset$, $\{1\}$, and $\{1, 2\}$.

First, Eq. (1a)=0 may or may not have real roots. If it does not have real roots, then □s□S^3, Eq. (1a)>0. If it has real roots s_1^{low} and s_1^{high}, and $\frac{1}{3} < s_1^{\text{low}} < s_1^{\text{high}} < 1$, then for the region s_1□$(s_1^{\text{low}}, s_1^{\text{high}})$, Eq. (1a)<0 and $\mathcal{L} = \emptyset$ is an equilibrium; Eq. (1a)>0 for otherwise.

Second Consider the relevant subspace of S^3 in which Eq. (1a)>0. It is obvious that Eq. (2b)=0 has two real roots of opposite signs when $F \geq \frac{\delta_2^2}{k}$. Thus, the relevant subspace can be partitioned into two parts by the curve Eq. (2b)=0. For the part Eq. (2b)<0, $\mathcal{L} = \{1\}$ is and equilibrium. For the other part Eq. (2b)>0, or equivalently, Eq. (5b)>0, it can be shown that as long as D_{12} is not too large, Eq. (5a)>0; thus $\mathcal{L} = \{1, 2\}$ is an equilibrium. This is proved as follows.

– Since Eq. (5b)=0 has two real roots of opposite signs: s_2^+ and s_2^-, Eq. (5b)>0 only for $s_2 \in (s_2^+, \min\{s_1, 1-s_1\})$.
– When $F \geq \frac{\delta_2^2}{k}$,

$$\frac{\partial(5a)}{\partial s_2} = -3F(k+s_2)^2 + 2(k+s_2)\left(\delta_1^2 - Fs_1\right) - 2ks_1\delta_1^2 < 0,$$

as long as $D_{12} < \frac{3(k+s_2)^2+2(k+s_2)s_1}{2k(k(1-s_1)+s_2)} = \widehat{D_f}$. This has to hold for $s_2 \in (s_2^+, \min\{s_1, 1-s_1\})$.
– It is straightforward that Eq. (5a)>(5b)>0 when $s_2 = s_1 < \frac{1}{2}$. Given the monotonicity of Eq. (5a) with respect to s_2 shown above, Eq. (5a)>0 for $s_2 \in (s_2^+, s_1)$.
– When $s_1 > \frac{1}{2}$, it can be shown that Eq. (5a)>(5b)>0 for $s_2 = 1b - bs_1$. To see this,

$$\begin{aligned}(5a)-(5b) = {} & \delta_1^2(k+1)\{(k+1)\left[4s_1^2-6s_1^2+4s_1-1(2s_1-1)k\right] \\ & +(1-2s_1)(k+s_1)^2\} \\ & +\delta_2^2(k+1)\{(k+1)[4s_1^3-6s_1^2+4s_1-1 \\ & +(2s_1-1)k]+(1-2s_1)(k+1-s_1)^2\}+\delta_3^2 \\ & \times(k+1)\{(k+1)\left[4s_1^3-6s_1^2+4s_1-1+(2s_1-1)k\right]\},\end{aligned}$$

and the coefficients on δ^2's are all positive. By monotonicity of Eq. (5a), (5a)>0 for $s_2 \in (s_2^+, 1-s_1)$.

It can be verified that, for k sufficiently large, $\hat{D}_d > \hat{D}_a > \min\{\hat{D}_b, \hat{D}_f\} > \min\{\hat{D}_c, \hat{D}_e\}$. Let $\hat{D}_{12} = \min\{\hat{D}_c, \hat{D}_e\}$, then $D_{12} \leq \hat{D}_{12}$ is sufficient for arguments (1)–(3) to hold. □

[26] Indeed, $s_3^+(s_2 = 0)(3c) \geq \frac{1}{3}$ is probably too strong. What we need precisely is $s_3^+(s_2)(3c) \geq \min\left\{s_2, \frac{1}{3}\right\}$.

References

Alesina, A.F., Rodric, D., 1994. Distribution politics and economic growth. Quarterly Journal of Economics 109, 465–490.

Altonji, J.G., Hayashi, F., Kotlikoff, L.J., 1992. Is the extended family altruistically linked? Direct tests using micro data. American Economic Review 82, 1177–1198.

Atkeson, A., Ogaki, M., 1996. Wealth-varying intertemporal elasticities of substitution: evidence from panel and aggregate data. Journal of Monetary Economics 38, 507–534.

Becker, G., Lewis, H.G., 1973. On the interaction between the quantity and quality of children. Journal of Political Economy 81, S279–S288.

Becker, G., Tomes, N., 1976. Child endowment and the quantity and quality of children. Journal of Political Economy 84, S143–S162.

Bénabou, R., 1996. Inequality and growth. In: Bernanke, B.S., Rotemberg, J.J. (Eds.), National Bureau of Economic Research Macro Annual, Volume 11. MIT Press, Cambridge, MA, pp. 11–74.

Bénabou, R., 2000. Unequal societies: income distribution and the social contract. American Economic Review 90, 96–129.

Bernheim, B.D., Whinston, M.D., 1986. Menu auctions, resource allocation, and economic influence. Quarterly Journal of Economics 101, 1–31.

Birdsall, N., Ross, D., Sabot, R., 1995. Inequality and growth reconsidered: lessons from east Asia. World Bank Economic Review 9, 477–508.

Bourguignon, F., 1981. Pareto-superiority of unegalitarian equilibria in Stiglitz' model of wealth distribution with convex savings function. Econometrica 49, 1469–1475.

Cameron, S.V., Heckman, J.J., 1998. Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males. Journal of Political Economy 106, 262–333.

Campante, F.R., Ferreira, F.H.G., 2007. Inefficient lobbying, populism, and oligarchy. Journal of Public Economics 91, 993–1021.

Cook, C.E., 1998. Lobbying for higher education: How colleges and universities influence federal policy. Vanderbilt University Press, Nashville, TN.

Deaton, A., Paxson, P., 1994. Intertemporal choice and consumption inequality. Journal of Political Economy 102, 437–467.

Deininger, K., Squire, L., 1996. A new data set measuring income inequality. World Bank Economic Review 10, 565–591.

Dixit, A.K., Grossman, G.M., Helpman, E., 1997. Common agency and coordination: general theory and an application to tax policy. Journal of Political Economy 105, 752–769.

Esteban, J., Ray, D., 2006. Inequality, lobbying and resource allocation. American Economic Review 96, 257–279.

Ferreira, F.H.G., 2001. Education for the masses? The interaction between wealth, educational and political inequalities. Economics of Transition 9, 533–552.

Filmer, D., Pritchett, L., 1999. The effect of household wealth on educational attainment: evidence from 35 countries. Population and Development Review 25, 85–120.

Filmer, D., Pritchett, L., 2001. Estimating wealth effects without expenditure data — or tears: an application to educational enrollments in states of India. Demography 38, 115–132.

Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. Review of Economic Studies 60, 35–52.

Glomm, G., Ravikumar, B., 1992. Public versus private investment in human capital: endogenous growth and income inequality. Journal of Political Economy 100, 818–834.

Green, L., Myerson, J., Lichtman, D., Rosen, S., Fry, A., 1996. Temporal discounting in choices between delayed rewards: the role of age and income. Psychology and Aging 11, 79–84.

Grossman, G.M., Helpman, E., 1994. Protection for sale. American Economic Review 84, 833–850.

Hallak, J., Poisson, M., 2002. Ethics and corruption in education. Policy Forum Series No. 15. International Institute for Educational Planning/UNESCO.

Hanushek, E.A., 1986. The economics of schooling: production and efficiency in public schools. Journal of Economic Literature 24, 1141–1177.

Hanushek, E.A., 1995. Interpreting recent research on schooling in developing countries. World Bank Research Observer 10, 227–246.

Kerckhof, A.C., Fogelman, K., Manlove, J., 1997. Staying ahead: the middle class and school reform in England and Wales. Sociology of Education 70, 19–35.

Lawrance, E.C., 1991. Poverty and the rate of time preference: evidence from panel data. Journal of Political Economy 99, 54–77.

Lee, W., Roemer, J.E., 1998. Income distribution, redistributive politics, and economic growth. Journal of Economic Growth 3, 217–240.

Lloyd-Ellis, H., 2000. Public education, occupational choice, and the growth–inequality relationship. International Economic Review 41, 171–201.

Meltzer, A.H., Richard, S.F., 1981. A rational theory of the size of government. Journal of Political Economy 89, 914–927.

Mincer, J., 1970. The distribution of labor incomes: a survey with special reference to the human captial approach. Journal of Economic Literature 8, 1–26.

Perotti, R., 1996. Growth, income distribution, and democracy: what the data say. Journal of Economic Growth 1, 149–187.

Persson, T., Tabellini, G., 1994. Is inequality harmful for growth? Theory and evidence. American Economic Review 84, 600–621.

Pollak, R.A., 1988. Tied transfers and paternalistic preferences. American Economic Review Papers and Proceedings 78, 240–244.

Rodriguez, F.C., Essays on the Political Economy of Inequality, Redistribution and Growth. Ph. D. thesis, Harvard University; 1998.

Saint-Paul, G., Verdier, T., 1993. Education, democracy and growth. Journal of Development Economics 42, 399–407.

Shavit, Y., Blossfeld, H., 1993. Persistent inequality: Changing educational attainment in thirteen countries. Westview, Boulder, CO.

Spencer, A.C., 1999. The new politics of higher education. In: Jacqueline, EK (Ed.), Financing a College Education: How It Works, How It Is Changing. American Council on Education and the Oryx Press1999, Phoenix, AZ, pp. 101–119.

Su, X., 2004. The allocation of public funds in a hierarchical educational system. Journal of Economic Dynamic and Control 28, 2485–2510.

Thrupp, M., 2001. Education policy and social class in England and New Zealand: an instructive comparison. Journal of Education Policy 16, 297–314.

UNESCO. Statistical Yearbook, various issues and URL: http://www.uis.unesco.org./url.

World Bank. World Development Indicator CD-ROM, 2004.