Working Papers / Documents de travail

Education Politics, Schooling Choice and Public School Quality: The Impact of Income Polarisation

Majda Benzidia
Michel Lubrano
Paolo Melindi-Ghidi

WP 2016 - Nr 42

Education Politics, Schooling Choice and Public School Quality: The Impact of Income Polarisation

Majda Benzidia* Michel Lubrano* Paolo Melindi-Ghidi*†

November 29, 2016

Abstract

Do communities with the same level of inequality but a different level of income polarisation perform differently in terms of public schooling? To answer this question, we extend the theoretical model of schooling choice and voting developed by de la Croix and Doepke (2009), introducing a more general income distribution characterised by a three-member mixture instead of a single uniform distribution. We show that not only income inequality, but also income polarisation, matters in explaining disparities in public education quality across communities. Public schooling is an important issue for the middle class, which is more inclined to pay higher taxes in return for better public schools. Contrastingly, poorer households may be less concerned about public education, while rich parents are more willing to opt-out of the public system, sending their children to private schools. Using micro-data covering 724 school districts of California and introducing a new measure of income polarisation, we find that school quality in low-income districts depends mainly on income polarisation, while in richer districts it depends mainly on income inequality.

JEL codes: I24, D31, D72, H52, C11.
Keywords: Schooling Choice, Income Polarisation, Probabilistic Voting, Education Politics, Bayesian Inference.

*Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS, EHESS and Centrale Marseille. Centre de La Vieille Charité, 2 rue de la Charité, 13002 Marseille, France.

†Corresponding author: paolo.melindi-ghidi@univ-amu.fr.

[0]**Acknowledgments:** This work was carried out thanks to the support of the A*MIDEX project (No ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR).

1 Introduction

On both sides of the Atlantic, the growth period corresponding to the golden sixties was accompanied by the growth of the middle class (see e.g. Levy and Murnane 1992). Essentially, the middle class was composed of educated white collar workers, ranging from permanent teachers to journalists and lawyers.[1] This upward social mobility was made possible by the emphasis on schooling and college education. The quality of schooling and consequently (local) public spending became a very important issue for the middle class, which saw the immediate social consequences for itself (Alesina and Glaeser 2004, Chap 7). This growth of the middle class was accompanied by a decrease in inequality till the early 1980s (Piketty and Saez 2003). The picture totally changed after that date. Income inequality greatly increased both in the US and in some European countries, mainly because the income of the upper deciles increased much more rapidly than that of the lower deciles. This movement was eventually accompanied by a reduced social mobility and a decline in social mobility. Booza et al. (2006) point out that the proportion of US metropolitan families earning middle incomes went down from 28 percent in 1970 to 22 percent in 2000. This decline in the size of the middle class, combined with a sharp rise in housing prices, led to an increase in urban polarisation. Middle-income neighbourhoods as a proportion of all metropolitan neighbourhoods declined from 58 percent to 41 percent over the same period. There was thus an amplification of inequality and polarisation, the middle class being obliged to move to poor neighbourhoods while the richer part of the population was concentrated in very rich areas.

Several papers have stressed the relationship between income inequality and public education quality. Card and Krueger (1992), using US data over 1920-1949, illustrate the influence of school quality on the rate of return on education. Based on the theory of the median voter, Meltzer and Richard (1981) show that higher inequality leads to more redistribution through higher taxation. More recently, this is corroborated for instance in Corcoran and Evans (2010) who, using a panel of US school districts, examine the relationship between income inequality and fiscal support for public education and find a positive relationship between inequality and public school spending. See also Soares (2003) who, in an overlapping generation model, shows that public funding for education is compatible with self-interest as it will increase the future income of the voters. However, it is not clear what happens when there are two competing systems of education, private schools and public schools. The proportion of private enrolment varies greatly over

[1]See for instance Mills (1951) for a tentative definition of the middle class.

countries, as do the system of financing for private schools. If financing of private schools is totally private, as in the US, the median voter theory would predict that public funding for public education is not likely to be affected. Yet de la Croix and Doepke (2009) propose a model with probabilistic voting (and an empirical test) where income inequality is the main determinant of segregation in the schooling system. Along these lines, Melindi-Ghidi (2016) shows that their result persists even though the model allows for the geographical mobility of households between school districts.

The idea we explore in this paper is that within a probabilistic voting context, inequality is not the only determinant of preferences for public funding of education, partly because income distribution cannot be represented by this indicator alone. While income polarisation was long accepted as representing the disappearance of the middle class by many sociologists after Max Weber (Andreski 2006, p. 105), proof that inequality was a different concept from polarisation was first given in Wolfson (1994) with axiomatisation provided by Esteban and Ray (1994). Actually, income inequality characterises the dispersion of an income distribution. It can be measured by various indices, such as the coefficient of variation or the Gini index. However, income polarisation characterises the increase of the ends of a distribution at the expense of its centre, moreover both its definition and its measurement are more complex. Polarisation can correspond to the decline of the middle class (Foster and Wolfson 2010) or to a distance between predefined groups (Esteban and Ray 1994). Polarisation is the source of potential conflicts (Esteban and Ray 1994) and leads to an uneven society, whereas a middle class consensus leads to more economic development and more investment in human capital as argued in Easterly (2001).

The model developed in our paper is able to take account of income inequality as well as the complex shape of income distribution. The main objective is to explain why we observe school districts with the same level of inequality, but with very different types of public education policies. Moreover, it is the first theoretical attempt in this literature to explain the effects of a major trend: *income polarisation*. The main theoretical contribution of the paper lies in analysing how differences in the income composition of school districts affect public education policies and school choice. We show that the effect of a shrinking middle class relative to the ends is ambiguous, and depends on the population composition of each school district. In particular, in districts populated by far more poor than rich households, income polarisation is more likely to negatively impact public schooling quality. This stems from the fact that the positive variation in the rate of participation in public education as a consequence of increasing polarisation is not offset by increased tax revenue to finance the public education system.

From a theoretical perspective, the model developed in this paper is an extension of the pioneering setting developed by de la Croix and Doepke (2009). Compared to their work, however we assume a more realistic income distribution which allow us to incorporate in their model a measure of income polarisation within school districts. While the original model is able to capture the effect of disparities in income inequality on public education policies, it is not able to determine whether these policies are also affected by changes in income distribution that do not alter inequality within districts. Put differently, if income distribution is represented with a standard uniform distribution, the impact of income polarisation on public education policies and schooling choices cannot be taken into account. Since the intended contribution of this paper is to assess the main consequences of income polarisation on public education policies, we introduce a parametric form for an income distribution which could be formulated to include parameters explaining this phenomenon. Thus, we propose a hybrid mixture model consisting of two uniform and one Pareto distributions.[2]

Using the ACS and ELSI databases, we aim at confirming this hypothesis through a two-regime regression model. We focus on determining whether polarisation impacts public spending and thus public schooling quality. We find that in school districts largely populated by poor households income polarisation matters, while when there is a majority of rich households, inequality is the major factor explaining public education spending.

The paper is organised as follows: section 2 provides background, with important stylised facts for the State of California. We also introduce an enriched income distribution that allows income polarisation to be considered in the analysis. In section 3 we extend the theoretical model of de la Croix and Doepke (2009) including the proposed income distribution. We examine the theoretical impact of income polarisation on public education policies and schooling choice. Section 4 focuses on the estimation of our enriched mixture distribution and indices. Section 5 presents an empirical analysis of the main theoretical results. Section 6 concludes.

2 From Inequality to Polarisation

Both the empirical literature, as well as its theoretical counterpart, highlight the effect of income inequality on public school quality. Inequality measures focus on the dispersion of a given income distribution. However, outside the lognormal framework, inequality alone does not adequately describe the

[2] To this end, we will need an income distribution composed of three income classes: the poor, the middle class and the rich.

shape of the income distribution, and polarisation becomes an important feature. This section attempts to detail the relationship between school quality and the shape of the income distribution, represented by both inequality and polarisation. Using California school district data for the year 2011-2012, we analyse the respective roles played by income polarisation and inequality. We first propose some stylised facts on public education quality and its relation to income. Second, we present a model for an income distribution more general than the uniform distribution considered in de la Croix and Doepke (2009) and the Pareto distribution considered in Arcalean and Schiopu (2016), but which allows for both analytical results and the empirical measurement of income polarisation.

2.1 The ACS and ELSI Databases

The American Community Survey (ACS)[3] is conducted every year. It provides demographic, social, economic, and housing data for the US. For educational data, we refer to the Elementary/Secondary Information System (ELSI)[4] which provides information on public and private schools.[5] We restrict our analysis to households with children enrolled in schools (public or private) for the year 2011-2012. These data sets are available by school district, grouped by state. We chose the state of California because it is the most densely populated state in the US. According to median household income, it is the third richest state, but according to income per capita, it is only 15th, very close to Illinois. There are rich counties in California, but there are much richer counties in Maryland for instance. California also has some very poor counties, mainly populated by Latin Americans or Asians. The state contains large cities with a high proportion of households living below the official poverty line (like Fresno, with a poverty rate of 31.5% in 2012). But there are also large, rich cities like San Francisco or San Jose, where the poverty rate is much lower (13% for San Francisco). The fertility rate, 1.90 in 2011, is just above the average. California is therefore a good choice because of its large size and the income diversity of its population.

After merging these two data bases, ACS and ELSI, our sample covers 724 school districts, documenting income distribution over the school district

[3] Available on: http://nces.ed.gov/programs/edge/demographicACS.aspx

[4] Available on: http://nces.ed.gov/programs/edge/tableViewer.aspx

[5] These two data bases are particularly interesting as they provide information by school district within a selected US state. Such information was previously available only through custom tabulations. These two data bases also make computation easier as they provide information not only for the total population but also for particular categories of children and parents.

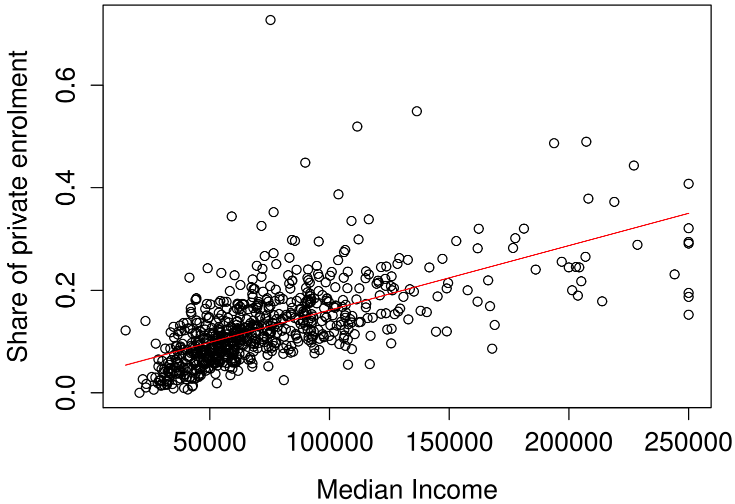


Figure 1: Relationship between share of private school enrolment and median household income for California in 2012

(grouped data for fixed classes of income), education (pupil/teacher ratio, instructional expenditure, numbers in private school enrolment...), federal and local tax revenues, fertility, and many other items relevant to our model. We now analyse some stylised facts concerning schooling choice, the influence of inequality *versus* polarisation on education quality. In the final section of this paper, we use this data set to empirically test the theoretical model of section 3.

2.2 Stylised Facts

When it comes to schooling, a decision many parents face is whether to send their children to private or to public schools. Different factors impact their decisions. In the US, public schools are financed through federal, state, and local taxes, while private school funding comes from a variety of sources, including tuition fees. These can be very high, whereas public schooling is offered free of charge. The ensuring gap between the wealthiest and poorest households, with only the wealthier able to provide private education for their children, means that this system can generate school segregation by income not only between but also within school districts.

Public schools are managed within school districts which are governed by elected councils, in a form of local government placed under the responsibility of the state (California in our case study). The elected council has the authority to manage local public schools. In the Californian school districts, we observe a clear income-related segregation at school district level. Figure 1 shows the empirical relationship between the share of enrolment in private schools and household median income, with the school district being the unit of observation. There is a strong positive (0.63) and highly significant correlation between income and the share of private school enrolment (nearly zero $p-$values for the Pearson correlation coefficient), which is one possible explanation for the variation in proportion of private schooling between school districts.

The data shows that the lowest rate of private school enrolment in Californian school districts is 0.006, which is very low, while the highest is 0.727, which is very high. The median rate is comparatively low, with 0.116. We do therefore observe school segregation by income in some districts, but it is not widespread. Within each school district, there is at least one household that chooses private schooling for its children and at least one that chooses public schooling. Presumably because school quality is comparatively high in California, our database contains no school district where only one alternative is chosen by all households.

Parents who can afford private schools may prefer to opt-out of the public school system because they have expectations of higher schooling quality in the private system. Schooling quality can be measured in different ways. The number of pupils per class has long been used as an indicator, starting with the Coleman report (1966), and is measured by the student-to-teacher ratio. More recent studies, such as de la Croix and Doepke (2009), prefer to use expenditure variables: schooling quality is measured by the level of instructional or total expenditure allocated to public schools. Why do some school districts benefit from greater financial means than others? As explained above, partly because of the level of local taxes, which is voted by the parents living in the school district and thus depends on the income distribution of that district. However, note also that the federal state can decide to compensate for lower local public funding in poorer school districts.

Although the literature contains many references to the relationship between income inequality and schooling quality, its nature is not clear, especially when public and private schools compete (see the papers cited in the introduction). Figure 2 provides two plots illustrating the relationship between public schooling quality and income inequality measured by a Gini coefficient. On the left side, we use public instructional expenditure as an indicator of public schooling quality, while on the right side, we use the public pupil-teacher ratio.[6] What we observe from our school district data for California is that inequality and schooling quality are totally unrelated. We get very low and non significant correlation coefficients with $p-$values of respectively 0.81 and 0.61.

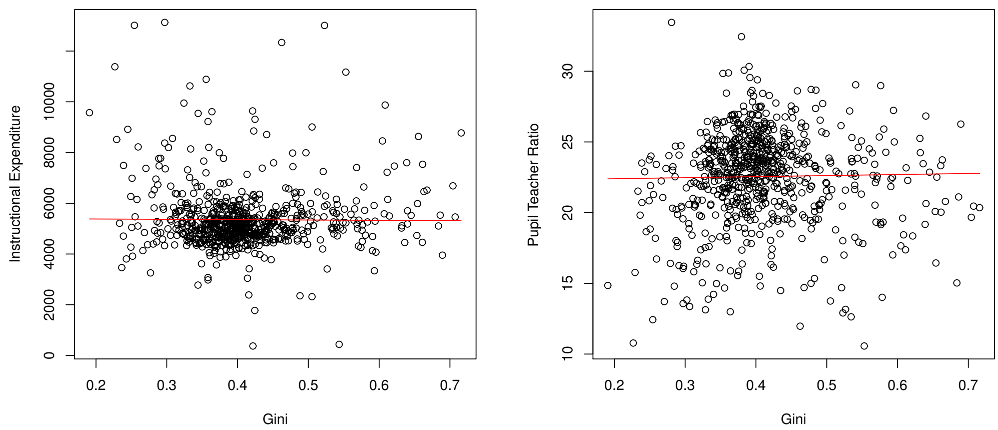
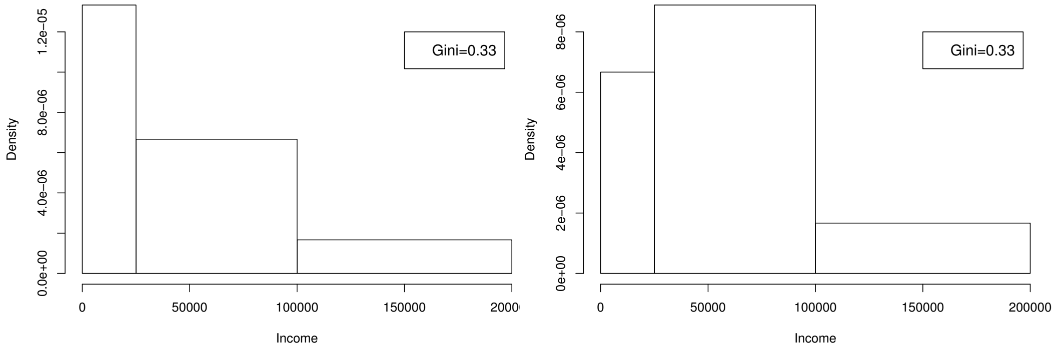


Figure 2: Relationship between public schooling quality and household income inequality.

However, focusing only on inequality may mean that we miss other important features of an income distribution. As an illustration, let us consider in Figure 3 two income distributions represented by a mixture of three uniform distributions and having the same Gini coefficient, but totally different proportions of middle class households. This figure illustrates the difference between polarisation and inequality, and explains why a model of income distribution that allows for both analytical results and the possibility of income polarisation is highly desirable. The more polarised a school district is, the higher the proportion of poor and rich compared to the middle class. This will have implications for public schooling, as we assume that the middle class tends to favour high public schooling quality.

By replacing the Gini index in Figure 2 with the measure of polarisation detailed below, we obtain Figure 4. Now the correlation between public education spending and our new indicator becomes strongly significant. On

[6]According to the National Center for Education Statistics, instructional expenditures are defined as "Expenditures for activities related to the interaction between teachers and students. Include salaries and benefits for teachers and teacher aides, textbooks, supplies and purchased services. These expenditures also include expenditures relating to extracurricular and co-curricular activities".



]

Figure 3: Two income distributions with the same level of inequality

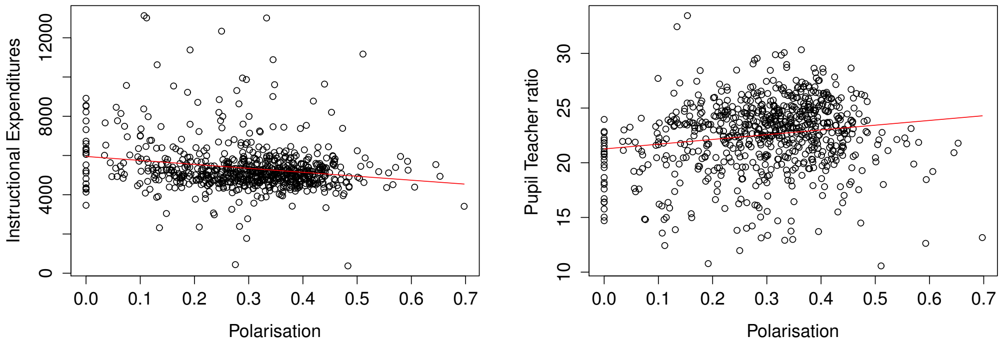


Figure 4: Relationship between public schooling quality and polarisation.

the left side of Figure 4, relating polarisation and instructional expenditure, the Pearson's product-moment correlation is -0.19, with a near zero p-value. On the right side of Figure 4 relating polarisation and the pupil-teacher ratio, the Pearson's product-moment correlation is 0.16 with a near zero p-value. We thus see a clear negative influence of polarisation on schooling quality when measured by both indicators.

2.3 An Enriched Income Distribution with Polarisation

In this section, we focus on a parametric form for an income distribution including in its formulation parameters directly monitoring polarisation. Polarisation is seen as the collapse of the middle class (Foster and Wolfson 2010), but also as a distance between predefined groups (Esteban and Ray 1994). Combining those two approaches, we propose an income distribution with three predefined income classes. To do so, we must specify two values x_1 and x_2 which are class boundaries. The poor are those with an income lower than x_1, the middle class those with an income between x_1 and x_2. The rich those with an income higher than x_2. Select of these values is discussed in section 4. Income is assumed to follow a uniform distribution inside each class. We thus have a mixture of three uniform distributions. If g is a parameter monitoring the size of the middle class and β a parameter monitoring the relative balance between the poor and the rich, we have:

$$\begin{aligned} f(x) &= \frac{g\beta}{x_1}\mathbb{1}(x < x_1) + \frac{1-g}{x_2 - x_1}\mathbb{1}(x_1 \leq x < x_2) \\ &+ \frac{g(1-\beta)}{x_{max} - x_2}\mathbb{1}(x_2 \leq x < x_{max}), \end{aligned} \tag{1}$$

where $\mathbb{1}(\cdot)$ is the indicator function. This distribution is not too restrictive, in view of the fact that we only have access to grouped data anyway. If the last income class is bounded, this solution is perfectly valid. Once the boundaries are fixed, the two parameters β and g are perfectly identified as $1 - g$ is equal to the proportion of households within the middle class boundaries and β can be recovered using the proportion of poor. However, when the top class of the grouped data is open, we have to consider a slightly different formulation, a mixture of two uniforms and of a Pareto:

$$\begin{aligned} f(x) &= \frac{g\beta}{x_1}\mathbb{1}(x < x_1) + \frac{1-g}{x_2 - x_1}\mathbb{1}(x_1 \leq x < x_2) \\ &+ g(1-\beta)\frac{\alpha x_2^\alpha}{x^{\alpha+1}}\mathbb{1}(x \geq x_2). \end{aligned} \tag{2}$$

Equipped with these two distributions, we can now explain the relationship between income polarisation and inequality.

2.4 Polarisation and Inequality in our Income Distribution Mixture

Let us now discuss polarisation in the context of our mixture model. The shape of the centre of the distribution, and consequently the size of the middle class, is monitored by the value of $g \in [0,1]$. Its relative size is maximum for $g = 0$ and minimum for $g = 1$, as evident from (2). The final shape of the distribution is monitored by the balance between the poor and the rich with β. For $\beta = 0$ we have no poor households, and for $\beta = 1$ we have no rich. For these two extreme cases, we cannot have polarisation, because one of the extreme groups disappears. For $\beta = 0.5$, the distance between the poor and the rich is maximum. Combining these two aspects, we propose a polarisation index between $[0,1]$ for this income distribution:

$$Pol = 4g\beta(1-\beta)).$$

This measure is maximum and equal to 1 when $g = 1$ (no middle class) and $\beta = 0.5$ (equal number of rich and poor). It is 0 when either $g = 0$ or when either the rich or the poor group disappears ($\beta = 0$ or $\beta = 1$).

Polarisation and inequality are nevertheless linked, despite being different in nature. In Figure 5, we plot our polarisation index against a Gini index for school district incomes. The Pearson's product-moment correlation is 0.29

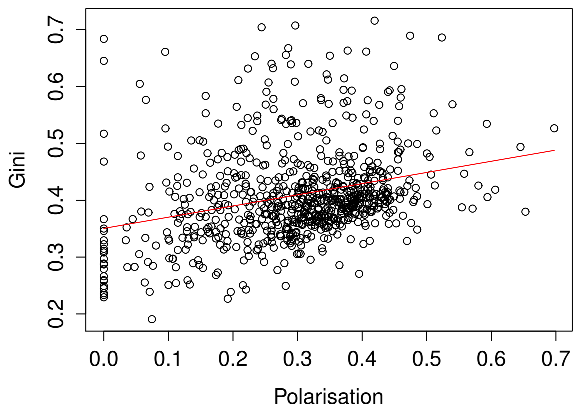


Figure 5: Relationship between polarisation and inequality in California

(and a near zero p-value) for this sample. However, the slope of regression line is far from 1. Moreover the plot is rather scattered. There are thus districts that have the same level of inequality but totally different levels of polarisation.

3 A Theoretical Model of Income Polarisation and Education Politics

The theoretical model analysed in this paper is an extension of the pioneering model without government commitment built by de la Croix and Doepke (2009). The authors develop a model of schooling choice and voting on public education, assuming that income is distributed according to a standard uniform distribution with bounded support $[1-\sigma, 1+\sigma]$. This assumption allows the authors to study the impact of inequality, proxied by the dispersion parameter σ, on the quality of the public school system. However, this simplifying assumption on income distribution does not allow us to describe the real distribution we actually observe within U.S. school districts. Nor can it explain why, in school districts with the same level of inequality, differentials in public schooling quality are observed.[7] The assumption of a uniform income distribution cannot account for the role that income polarisation might play in explaining disparities in schooling quality. For these reasons, we assume that the distribution of income is characterised by a mixture of two uniform distributions and of a Pareto distribution. Our objective is to understand the main effects that an income distribution accounting for differences between and within social classes might have on public policies, and therefore on schooling quality and segregation.

3.1 Household Decision Problem: The Theoretical Setup of de la Croix and Doepke (2009)

Our theoretical starting point is the problem of a representative household as developed in de la Croix and Doepke (2009). It is a model with endogenous fertility in which households choose their consumption level c_t, decide their number of children n_t, and whether to educate them in public or private schools. Public education is free of charge, while private education involves a tuition fee.[8] When fertility and education decisions are taken, households vote for a rate of income tax rate to finance public education spending.[9] The representative agent is endowed with an additive and separable utility function where $\gamma \in \mathbb{R}^+$ is the overall weight attached to children and $\eta \in$

[7] In section 4, we estimate the California income distribution showing that this mixture is a good representation of the analysed date. See Figure 6.

[8] For simplicity, the unitary cost of private education is normalised to one.

[9] See de la Croix and Doepke (2009) for more details on the main theoretical assumptions and timing of the events. We replicate in this sub-section the household problem developed by de la Croix and Doepke (2009) to help the reader follow the theoretical extension proposed in our paper.

$(0, 1)$ is the relative weight of human capital quality:

$$u = \ln(c) + \gamma[\ln(n) + \eta \ln(h)]. \tag{3}$$

In this equation, $h = \max\{s, e\}$ represents the level of child human capital, i.e., the quality of education acquired by each child with s representing the quality of public schooling, proxied by public spending per child, while e represents private investment on education. Public and private education are mutually exclusive. Private education spending being assumed tax deductible, the budget constraint is simplified and writes:

$$c = (1 - \tau)[x(1 - \phi\, n) - n\, e], \tag{4}$$

where τ is the income tax rate, x the exogenous wage rate and ϕ the proportion of time allocated to raising one child, $(1 - \phi\, n)$ representing labour supply.[10] Maximising utility (3) under the budget constraint (4), it is possible to derive the desired number of children n and the optimal education investment e, for each choice of schooling type:

$$\begin{array}{lll} \textit{Public}: & e^s = 0 & n^s = \dfrac{\gamma}{\phi\,(1+\gamma)} \\ \textit{Private}: & e^e = \dfrac{x\eta\phi}{1-\eta} & n^e = \dfrac{\gamma(1-\eta)}{\phi(1+\gamma)}, \end{array} \tag{5}$$

implying $n^s > n^e$: parents choosing the public system have more children.[11]

We define the indirect utility function V^s, corresponding to choosing public education, and V^e, corresponding to choosing private education, by replacing the budget constraint (4) and the optimal decisions (5) in the utility function (3). The final schooling choice is made by comparing the two indirect utility functions V^s and V^e. The possibility that $V^e > V^s$ depends on the expected quality of public schooling $E[s]$ and only arises if the agent has an income greater than a threshold given by:

$$x > \tilde{x} \equiv \frac{E[s](1-\eta)^{\frac{\eta-1}{\eta}}}{\phi\eta}, \tag{6}$$

which is Lemma 2 in de la Croix and Doepke (2009). At a given wage x, the higher the expected quality of public schooling $E[s]$, the lower the probability of opting-out of the public education system. Since households have perfect foresight over the outcome of the political process and, consequently, over the policies adopted by the government, $E[s] = s$.

[10] There is an implicit constraint in the model so that the maximum number of children n is bounded by $1/\phi$.

[11] Since $e^s = 0$, for simplicity we will define $e^e = e$ in the paper.

3.2 Introducing Income Polarisation

Let us now consider our enriched income distribution (2). We first determine the participation rate in the public school system Ψ as the integral of this income distribution between 0 and the predetermined threshold $\tilde{x}$:

$$\Psi = \int_0^{\tilde{x}} f(x)dx = g\beta\frac{\tilde{x}}{x_1} + (1-g)\frac{\tilde{x}}{x_2 - x_1} - (1-\beta)g\left(\frac{\tilde{x}}{x_2}\right)^{-\alpha}. \qquad (7)$$

Following de la Croix and Doepke (2009), we assume that fertility and schooling choices are determined before the political process takes place. Consequently, the opting-out threshold $\tilde{x}$ can be taken as given. Therefore, the derivative of (7) with respect to β under the perfect foresight assumption is always positive and we have:

Lemma 1 *Given g, an increase in β the relative proportion of poor households compared to rich households, positively affects the participation rate in public education.*

For a proof, see Appendix A.1.

The effect of g on the participation rate is however ambiguous and depends on the relative position of the opting-out threshold $\tilde{x}$ with respect to the exogenous thresholds x_1 and x_2:

Lemma 2 *Given β, an increase in g, the share of the ends compared to the middle class has an ambiguous effect on the participation rate in public education:*

$$\frac{\partial\Psi}{\partial g} = \begin{cases} > 0 & if \quad \beta > \hat{\beta} \\ < 0 & if \quad \beta < \hat{\beta} \end{cases}$$

with

$$\hat{\beta} \equiv \frac{x_1\left((x_2 - x_1)^{-1} + x_2^{\alpha}\tilde{x}^{-(1+\alpha)}\right)}{1 + x_1 x_2^{\alpha}\tilde{x}^{-(1+\alpha)}}.$$

For a proof, see Appendix A.2.

Lemma 2 indicates that if the proportion of poor households in the non-middle class group is sufficiently high (low), i.e. $\beta > \hat{\beta}$ ($\beta < \hat{\beta}$), then an increase in parameter g positively (negatively) impacts the participation rate in the public education system. This result is quite intuitive and can be explained by construction, i.e., because education is a normal good. The rich can afford it and the poor can't. In other words, the income polarisation of each school district matters, even though income inequality at district level is the same.

3.3 Equilibrium and Income Polarisation under Probabilistic Voting

We now study how the voted policies in the theoretical model of de la Croix and Doepke (2009) are modified when we move from the simple uniform income distribution between $1-\sigma$ and $1+\sigma$ to the mixture model (2) defined over the support $[0, +\infty[$. This generalisation allows us to introduce a measure of income polarisation, whereas the original model can only account for income inequality. The analysis of the effects of the two parameters, β and g, on the political mechanism and therefore on the equilibrium is the main theoretical contribution of this section.

Each school district must have a balanced budget. The total spending for public schools, which is given by:

$$\int_0^{\tilde{x}} s\, n^s f(x)dx, \tag{8}$$

has to be equal to the total local income tax revenue. As both types of households, those sending their children to public schools and those sending their children to private schools pay taxes, the local tax revenue is:

$$\tau \int_0^{\tilde{x}} [x(1-\phi n^s)] f(x)dx + \tau \int_{\tilde{x}}^{\infty} [x(1-\phi n^e) - en^e] f(x)dx. \tag{9}$$

Since education spending is assumed tax deductible and fertility is endogenous, taxable income is the same whether parents choose public or private education. Indeed, using (5), it is easy to verify that $x\,(1-\phi\, n^s) = x(1-\phi\, n^e) - en^e \equiv x/(1+\gamma)$. We can rewrite the balanced budget rule of the local government as follows:

$$\frac{s\gamma}{(1+\gamma)\phi} \int_0^{\tilde{x}} f(x)dx = \frac{\tau}{1+\gamma} \int_0^{\tilde{x}} xf(x)dx + \frac{\tau}{1+\gamma} \int_{\tilde{x}}^{\infty} xf(x)dx. \tag{10}$$

Solving (10) and following Arcalean and Schiopu (2016), we are able to rewrite the government budget constraint so as to express the quality of public schooling as a function of the tax rate τ, the participation rate Ψ, and the mean of the income distribution μ:

$$s[\tau, \Psi, \mu] = \frac{\mu\tau\phi}{\Psi\gamma} \tag{11}$$

with

$$\mu \equiv \int_0^\infty x f(x)dx = g\beta\frac{x_1}{2} + (1-g)\frac{x_1+x_2}{2} + (1+\beta)g\frac{x_2\alpha}{\alpha-1}.$$

As already observed in Persson and Tabellini (2002) and de la Croix and Doepke (2009), the equilibrium choice under probabilistic voting is equivalent to maximising a weighted sum of the indirect utilities of individuals:

$$\Omega[\tau] = \int_0^{\tilde{x}} V^s[x, n^s, 0, s, \tau]f(x)dx + \int_{\tilde{x}}^\infty V^e[x, n^e, e, 0, \tau_t]f(x)dx. \qquad (12)$$

Using (3), (4) and (5) in order to implicitly define the two indirect utility functions V^s and V^e, we can rewrite (12) as follows:

$$\Omega[\tau] = \int_0^{\tilde{x}} \left(\ln\left[\frac{x(1-\tau)}{1+\gamma}\right] + \gamma\ln\left[\frac{\gamma}{\phi(1+\gamma)}\right] + \gamma\eta\ln[s[\tau,\Psi,\mu]]\right) f(x)dx$$
$$+ \int_{\tilde{x}}^\infty \left(\ln\left[\frac{x(1-\tau)}{1+\gamma}\right] + \gamma\ln\left[\frac{\gamma(1-\eta)}{\phi(1+\gamma)}\right] + \gamma\eta\ln\left[\frac{x\eta\phi}{1-\eta}\right]\right) f(x)dx. \qquad (13)$$

Using the government budget constraint (11), after some algebraical manipulations, the above social welfare function writes:

$$\begin{aligned}\Omega[\tau] &= \ln\left[\frac{1-\tau}{1+\gamma}\right] + \gamma\ln\left[\frac{\gamma}{\phi(1+\gamma)}\right] + \gamma\eta\ln\left[\frac{\mu\tau\phi}{\Psi_t\gamma}\right]\int_0^{\tilde{x}} f(x)dx \\ &+ \int_0^\infty \ln[x]f(x)dx \\ &+ \int_{\tilde{x}}^\infty \left(\gamma\ln[1-\eta] + \gamma\eta\ln\left[\frac{x\eta\phi}{1-\eta}\right]\right) f(x)dx]. \qquad (14)\end{aligned}$$

Now, taking the first-order condition with respect to τ for a maximum, we can express the voted tax rate in terms of participation rate in the public education system:

$$\tau = \frac{\gamma\eta\Psi}{1+\gamma\eta\Psi} \equiv \tau[\Psi], \qquad (15)$$

as well as the expected level of public education spending in terms of participation rate and mean income:

$$s = \frac{\mu\eta\phi}{1+\gamma\eta\Psi} \equiv s[\Psi,\mu]. \qquad (16)$$

First of all, note that the voted tax rate is an increasing function of the participation rate in public school while the public spending per student is a decreasing function. This result is quite intuitive and comes from the fact that in presence of private schools high-income parents might opt-out from the public education system. In contrast to the existing theoretical literature, since Ψ and μ are both functions of parameters β and g, our model is able to capture the direct impact of income polarisation on public policies and schooling choice. More precisely, the population composition of each school district may have a crucial impact on public policies even in school districts whose income inequality are similar. Put differently, our model is able to explain the empirical case in which disparities in public schooling quality emerge even in school districts where income inequality is at a similar level.

Note that, since agents are rational, taxable income does not depend on the participation rate and the perfect foresight condition on expected schooling quality holds. Therefore, the main results of de la Croix and Doepke (2009) on the existence and uniqueness of the equilibrium apply (see their proposition 1, p. 605). This statement is easily proved by following the analytical proof of proposition 1 in Arcalean and Schiopu (2016).

It should be noted that contrary to de la Croix and Doepke (2009), in our model the equilibrium level of schooling quality (16) depends on both the participation rate in public education and the mean of the income distribution. In particular, the higher the average income in the economy, the higher the public schooling spending per child $s[\Psi, \mu]$. This result is not surprising since it is reasonable to expect richer school districts to perform better in terms of public schooling than poorer school districts. However, a model with a standard uniform income distribution is not able to explain the plausible scenario in which, given parameters ϕ, γ, η, two school districts with the same level of inequality could perform differently in terms of schooling quality.

On this important theoretical and empirical issue, it is crucial to underline the main differences between our setting and the original model of de la Croix and Doepke (2009). First, given the opting-out threshold $\tilde{x}$, in our paper the fraction of households that choose public schooling does not solely depend on the dispersion of the income distribution. In our model, the participation rate in the public education system also depends on income polarisation and, therefore, on the internal composition of the population of each school district. If inequality reaches the same level across school districts, our theoretical set-up does not ensure that the voted public policies in each district will be the same. In other words, the assumed income distribution allows us to explain why public school spending and schooling quality could be different in economies with the same level of inequality.

A second crucial difference is that, in our set-up, neither the fully private regime nor the fully public regime can be an equilibrium outcome. Assuming a more realistic income distribution function over the support $[0, \infty]$ necessarily implies a certain level of schooling segregation by income. This result is in line with the empirical observations in section 2: in each school district we always observe a positive, even though very low, private school enrolment rate.

Third, our model is able to explain disparities in schooling quality through the channel of income polarisation. This is in line with our view that it is not only income inequality that matters but also how the income is distributed. Actually, if we consider as given the proportion of middle income households, the relative balance between poor and rich households can have important effects on public policies such as public education spending. We have the following lemma:

Lemma 3 *Given g, an increase in the relative proportion of poor β leads to a higher tax rate and lower public education spending:*

$$\frac{\partial \tau[\Psi]}{\partial \beta} > 0, \quad \frac{\partial s[\Psi, \mu]}{\partial \beta} < 0.$$

For a proof, see Appendix A.3.

Lemma 3 establishes that given the size of the middle income group, a school district with a larger share of poor than of rich households has a higher participation rate in public schooling, a higher tax rate and lower education spending per child. This theoretical result confirms the empirical observation that poor urban areas are characterised by public schools of low quality but a higher participation rate in the public education system.

We now concentrate on the crucial relationship between income polarisation - measured as the decline of the middle class - and public policies.

Proposition 1 *Given β, the effect of an increase in income polarisation on public education spending is negative (positive) if $\beta > max\{\hat{\beta}, \bar{\beta}\}$ ($\beta < min\{\hat{\beta}, \bar{\beta}\}$) and ambiguous if $\beta \in [min\{\hat{\beta}, \bar{\beta}\}, max\{\hat{\beta}, \bar{\beta}\}]$.*

For a proof, see Appendix A.4.

Proposition 1 give us a first important clue: If we want to analyse the effect of income polarisation, proxied by the disappearance of the middle class, the income distribution of the school district matters. The effect of a variation in the size of the middle class compared to the ends mainly depends on the relative balance between poor and rich households within each school district. More precisely, our model predicts that if the district contains a sufficiently large proportion of poor compared to rich households,

$\beta > \max\{\hat{\beta}, \bar{\beta}\}$, the relative reduction in the middle income group will negatively impact public education spending per child. In fact, in this particular scenario, an increase in the share of the 'ends' compared to the middle will positively impact the participation rate in public schooling. Therefore, even if the tax rate increases, our model predicts that in school districts populated by more poor than rich households, the effect of income polarisation on public education spending will be negative.

Conversely, when school districts have a larger proportion of rich than of poor households, $\beta < \min\{\hat{\beta}, \bar{\beta}\}$, the public education participation rate goes down as a consequence of increased income polarisation. The rich parents are more willing to invest in education and enrol their children in private schools. In this scenario, reduction in the size of the middle income group negatively impacts the public education participation rate and the voted tax rate. Therefore, as more households opt-out of the public schooling system, public spending per student increases.

However, the effect of income polarisation on public education spending is found to be ambiguous when the share of poor and rich households is relatively balanced, that is when $\beta \in [min\{\hat{\beta}, \bar{\beta}\}, max\{\hat{\beta}, \bar{\beta}\}]$. The effect will depend on the values of the exogenous parameters, on the relative position of the opting-out threshold $\tilde{x}$, as well as on the impact on the income distribution. More precisely, it will depend on the characteristics of each school district and on the relative variation in the public education participation rate with respect to the variation in the tax rate as a consequence of income polarisation.

Corollary 1 *The poorer the school district, the more likely income polarisation is to negatively impact public education spending.*

For an intuitive understanding of Corollary 1, consider the following. Given β, we observe that the sign of $\frac{\partial s[pol,\beta]}{\partial pol} > 0$ if and only if $\frac{\partial \mu}{\partial pol} > k \frac{\partial \psi}{\partial pol}$ with $k = \frac{\gamma \eta \mu}{1+\gamma \eta \psi}$. The larger μ, the larger the value of k. This implies that for a given public education participation rate, the effect of income polarisation on public school spending is more likely to be positive in rich school districts characterised by a high mean income. Moreover, since high-income parents are more willing to enrol their children in private schools, the public education participation rate will be lower in richer school districts. The opportunity to opt-out of the public education system, the fact that education is a normal good, and the political arguments leading poor and rich households to vote against redistribution, explain why we find that income polarisation has a negative (positive) impact on the quality of public schooling in poor (rich) school districts. In the next two sections, we will whether micro-data for California confirm our theoretical conclusions.

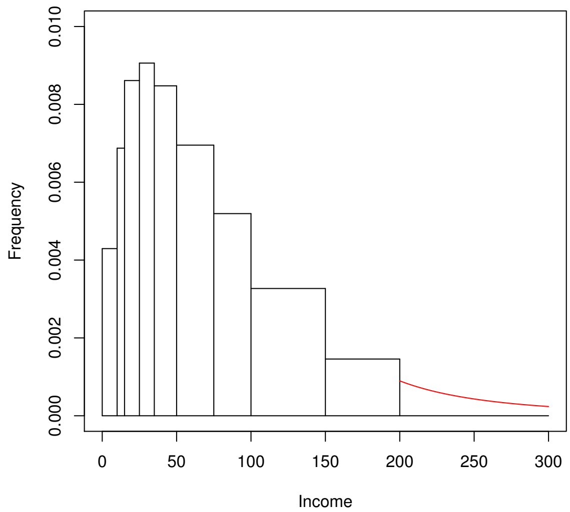


Figure 6: Household income distribution for California

4 Estimation of California's Income Distribution

In the ACS data base, information on household income distribution (with children in school) is provided at school district level grouped in ten unequal classes, with top-coding for the highest. The lowest class is households with a yearly income plus benefits of below \$10 000, while the highest class is households with a yearly income plus benefits above \$200 000. It is not possible to apply an equivalence scale for these income data, because we combined observations for both income and family composition.[12] We show this distribution for the whole state of California in Figure 6, including all school districts. The top open class is represented by a Pareto distribution. The way the Pareto parameter ($\hat{\alpha} = 2.28$) is estimated is explained below.

[12]This should not have much incidence because we restrict our attention to households with children and the average household size does not vary too much over school districts. On average, even if this is not indicated in our data sources, the figures seem to correspond to a household with two adults and two children, according to the new OECD equivalence scale.

4.1 A Stylised Income Distribution

We divide the ten classes into three groups representing the poor, the middle-class and the rich, so as to obtain an empirical model that matches our theoretical model. We represent the poor households as having a yearly income lower than \$25 000. In 2012, the official median household income was \$58 328 for California (Noss 2013, American Community Survey Briefs), while the median income of our sample (households with children) is \$63 477.[13] \$25 391 represents 40% of our median income while the next income class (\$35 000) represents 55% of that median income. It should be noted, however, that there is no relative poverty line in the US. Using \$25 000 as a poverty line, Table 1 shows 16% poor since the US Census Bureau gives a poverty rate of 17% for California in 2012, we are pretty safe. We now need

Table 1: Income distribution in Californian school districts (all California)

Classes	$x < 25$	$25 < x < 100$	$x > 100$
Percent	16	52	32
Number	621 264	1 982 006	1 198 231

to define the line that separates what we call the middle class from the richer part of the population. There is no universal definition of the middle class. Following Piketty and Saez (2003), the top decile of an income distribution represents both the upper middle class and the very rich. However, among the various of definitions of the middle class reported in Renwick and Short (2014), those referring to quantiles use the range between 0.25 and 0.75 of the income distribution. There are also definitions in terms of median income, with a lower bound between 0.50 and 0.75 of the median and an upper bound between 1.25 and 2.00 times the median. 150% of the median income would correspond in our case to an upper bound of \$95 216, and we take \$100 000 as the upper bound for the middle class. With this definition of middle class (between \$25 000 and \$100 000 a year), we have a middle class representing 52% of our sample. 32% of the households have an income greater than \$100 000.

This stylised income distribution is shown in Figure 7. It is interesting to compare the income distribution of each school district with that reference distribution, bearing in mind that the total number of rich is twice the number of poor. There are 331 districts out of 724 where the middle class is

[13] This figure is obtained by taking the median value of the reported median income of each of the 724 school districts.

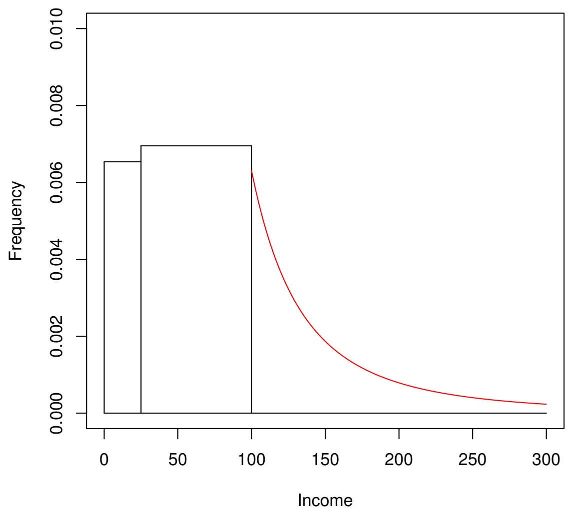


Figure 7: Stylised income distribution for California

dominant (more than 50%) and 125 districts where the rich are dominant. There is no district where the poor are dominant, but there are 169 districts where the poverty rate is greater than our average poverty rate of 16%.

4.2 Parameter Estimation

Our model for the income distribution has three parameters to be estimated for each school district. The poor and the middle class are represented by a uniform density, while the rich class is represented by a Pareto. There are 113 school districts where the top income class (income greater than \$200 000) is empty. In this case, we have chosen to represent the distribution of the rich by a simple uniform between finite bounds. Let us call n_j the number of households in income group j with $j = 1, \cdots, 3$ and $n = \sum_{j=1}^{3} n_j$ the total number of households in a school district. For each school district, we estimate first:

$$\hat{g} = 1 - n_2/n, \tag{17}$$
$$\hat{\beta} = n_1/(n\,\hat{g}), \tag{18}$$

using the ten income classes provided in our sample. The median value of the 724 values of $(1 - \hat{g})$ is 0.56 with a standard deviation of 0.15, which means that there is a majority of school districts where the middle class dominates.

The median of the 724 values of β (the relative proportion of poor) is 0.34 with a standard deviation of 0.26. The relative proportions of rich and poor households vary extensively.

Let us now turn to the estimation of α for the Pareto member. We start by defining nc_i as the number of households in each of the original ten income classes while nc_{10} represents the number of household in the top open income class (with an income greater than $x_{10} = \$200\ 000$):

$$\hat{\alpha} = \frac{\log(nc_8 + nc_9 + nc_{10}) - \log(nc_{10})}{\log(x_{10}) - \log(x_8)}, \tag{19}$$

where nc_8, nc_9 and nc_{10} represent the number of households in classes 8, 9 and 10 and x_8 the lower bound of class 8. Formula (19) is an adaptation of the formulae of Quandt (1966) (see also von Hippel et al., 2015) where the group representing the rich is obtained by aggregating three classes: 8, 9 and 10.

This formula does not work when the upper class nc_{10} contains more households than nc_9. This case happens in 57 school districts where the predetermined ten class boundaries do not adequately represent the right tail of the income distribution. In these cases, the top class has to be divided further into two income classes: say incomes between \$200 000 and \$300 000 for the tenth class and incomes greater than \$300 000 for a hypothetical eleventh class. We assume that the number of households in the newly defined tenth class is a fraction of the old class, for instance $\tilde{n}c_{10} = nc_{10}/1.4$, so that the top open class contains $\tilde{n}c_{11} = (1 - 1/1.4)nc_{10}$. With this extension, we estimate α as:

$$\hat{\alpha} = \frac{\log(nc_8 + nc_9 + \tilde{n}c_{10} + \tilde{n}c_{11}) - \log(\tilde{n}c_{11})}{\log(x_{11}) - \log(x_8)}. \tag{20}$$

We were able to estimate α in 611 cases out of 724.[14] However, with this simple method, we have 12 districts where $\alpha \leq 1$, meaning that we cannot calculate for our complete distribution. Where available we can use the mean in each school district to obtain an improved estimator for α which fulfils the constraint of $\alpha > 1$. The mean of our overall distribution is the weighted sum of the mean of each member:

$$\mu(\alpha, g, \beta) = g\,\beta\,\frac{x_1}{2} + (1-g)\,\frac{x_2 - x_1}{2} + g\,(1-\beta)\,\frac{\alpha\,x_2}{\alpha - 1},$$

where $x_1 = \$25\ 000$, $x_2 = \$100\ 000$ and $\alpha > 1$. Our previous estimate of α together with $\hat{g}$ and β provide an estimation for the empirical mean which

[14]The remaining cases correspond to the situation where $nc_{10} = 0$ and is represented by a mixture of three uniforms.

can be compared with the mean provided in the data for each school district. We can improve our first estimator of α by minimising in α the following loss function

$$(\mu(\alpha, \hat{g}, \hat{\beta}) - ms)^2 + (\alpha - \hat{\alpha})^2,$$

where $\hat{\alpha}$ is the initial estimator resulting from the application of (19) or (20) and ms is the empirical mean in the data set. The median value of the estimated α is 2.78 with a standard deviation of 1.30 and the minimum value is now 1.20.

4.3 Estimation of the overall Gini coefficient

When the procedure for computing a Gini coefficient with grouped data inspired from Gastwirth (1972) and Schader and Schmid (1994) is applied to our 724 school districts, the median value of the Gini coefficient is 0.40. This indicates rather high overall income inequality in all the school districts. However, the school districts are very heterogeneous, as the Gini ranges from 0.19 to 0.72. When we estimate our polarisation coefficient, we find even greater dispersion, with the coefficient ranging from 0.00 to 0.70. Figure 5 shows that the relationship between income inequality and income polarisation is rather weak for the 724 school districts and that inequality and polarisation represent two distinct notions.

5 Testing our Theoretical Model on Californian Data

Each of the 724 school districts is now equipped with an estimated income distribution, with its parameters β, g andα (when there is a Pareto member), a polarisation index and a Gini index. We have all the necessary ingredients to compare our theoretical model with the data. Essentially, the econometric model we use is a two-regime switching model where the change of regime depends on a threshold. There are strong arguments for adopting a Bayesian approach to make inference in this model, as argued in Bauwens et al. (1999, Chap. 8). Essentially, the distribution of the threshold parameter is non-standard and can be multi-modal. This type of situation does not lend itself to classical inference: neither for estimation, because we are never sure which maximum is found, nor for testing, because the asymptotic distribution of the threshold parameter is not standard, as detailed in Hansen (2000). With a Bayesian approach, we simply have to integrate over a given a priori range the posterior density of the threshold parameter; and for integration, multi-modality is of no practical importance. We merely need to choose the range of integration with care because it conditions the identifiability of the model. This is the only informative prior information that we have to provide. Finally, as Bayesian inference provides the small sample posterior density for each parameter, testing is not a problem.

5.1 Determining the Opting-Out Threshold

We first have to determine the income threshold above which there is a majority of private school enrolment (the opting-out threshold). Is private schooling reserved for the rich households alone, or is part of the middle class also concerned? First of all, from Figure 8, we see that public schools play a very important role in California, because there is no district where the public enrolment is zero and the 1% quantile is equal to $\Psi = 0.60$. We need to bear in mind that the data represent school districts and not households, so these are average figures over a district.

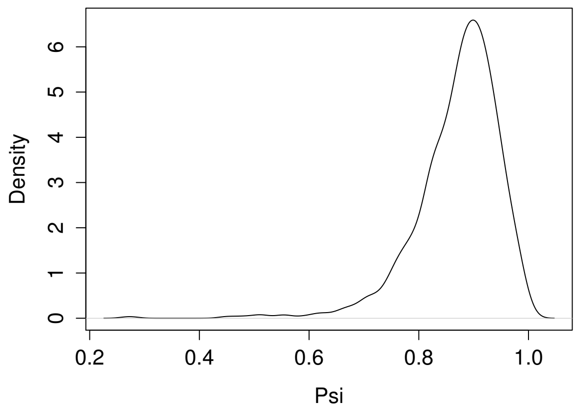


Figure 8: Rate of public school enrolment

We propose a two-regime regression model explaining the rate of public school enrolment Ψ where the threshold between the two regressions is the level of mean income divided by the average number of children per household. Dividing mean income by 1 000, we choose the range [60,150] as prior information on the threshold. Our prior range covers both the middle class and the richer class. The posterior results we get are as follows:[15]

$$\begin{aligned}\Psi &= \underset{(66.3)}{0.861} - \underset{(-2.25)}{0.055}g + \underset{(11.70)}{0.118}\beta \quad \text{for } x < \tilde{x} \\ \Psi &= \underset{(26.5)}{0.899} - \underset{(-4.45)}{0.182}g - \underset{(-0.09)}{0.007}\beta \quad \text{for } x \geq \tilde{x}\end{aligned}$$

The posterior mean of the opting-out threshold is $\tilde{x} = \$106\ 187 \times 1.62$ (with standard deviation equal to $\$1\ 925 \times 1.62$) where 1.62 is the mean number of children per household in the sample. The posterior mean of the variance of the error term of the regression model is $\sigma^2 = 0.00333$. There are 611 observations in the first regime (poor), with an average polarisation index of 0.311, a mean Gini of 0.394 and an average public school enrolment of 0.89, a mean household income of \$80 448. There are 113 observations in the second regime (rich), with a mean polarisation index of 0.224, a mean Gini of 0.490 and an average public school enrolment of 0.77, a mean household income of \$356 201.

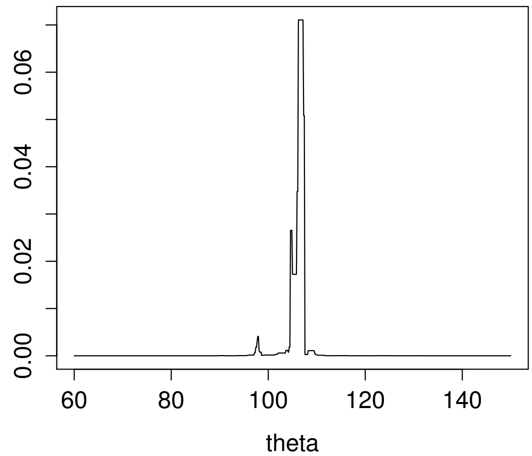


Figure 9: Posterior density of the opting-out threshold $\tilde{x}$

The opting-out threshold is relatively high involving on average only the richer class. Note that this does not mean that there is no member of the lower classes choosing the private sector. Polarisation is higher and inequality lower in the first regime, while inequality is higher and polarisation lower in

[15]Student ratios are given between parentheses. Average characteristics of the two subsamples were computed as a byproduct of integration. The reader used to significance codes with stars can apply the following conversion scale: *** $P \leq 0.001$, ** $P \leq 0.01$, * $P \leq 0.05$ correspond to a Student ratio greater than 3.29, 2.58 and 1.96 respectively.

the richer group. Figure 9, which displays the posterior density of the opting-out threshold (divided by the average number of children), shows that the latter is rather well determined.

5.2 Explaining Public Schooling Quality

Our Proposition 1 says that, depending on the level of a given β (the relative proportion of poor), public education spending depends on polarisation. This contrasts with de la Croix and Doepke (2009) according to whom the level of public education spending is on the whole negatively affected by inequality. Our theoretical results suggest an econometric model with several regimes, where the change of regime is determined by the value of β. For the dependent variable, we have the choice between *Instructional Expenditure per Pupil* and *Total Expenditure per Pupil* which covers a broader spectrum of expenditure.[16] These expenditures are financed by tax revenues, essentially from three sources. The main source should be local taxes, which are property taxes. In addition, the State of California provides state tax revenue. Finally, the Federal Government provides tax revenue which acts as a form of redistribution. The sum of these three tax revenues should cover our *Total Expenditure per Pupil*. We shall focus on the effect of local and federal tax revenues.

Following Proposition 1, we expect that, for low values of β, say $\beta < b_1 = \min(\hat{\beta}, \bar{\beta})$, polarisation should have a positive effect on public education spending, while for high values of β, say $\beta > b_2 = \max(\hat{\beta}, \bar{\beta})$, polarisation should have a negative sign. For intermediate values ($b_1 < \beta < b_2$), our theory does not provide a precise sign indication. Therefore, the correct econometric model would be a three-regime model explaining the log of public education spending by our polarisation index, a Gini index (in order to have a point of comparison with de la Croix and Doepke (2009)) and the log of

[16] The web site of the National Center for Education Statistics provides the necessary definitions. *Instructional Expenditure per Pupil* covers mainly wages and activities related to the interaction between teachers and students. *Total Expenditure per Pupil* includes the previous expenditure and adds maintenance, investment, interest payments, student support, food, administration, etc. There is one abnormal value for the *Instructional Expenditures per Pupil* at $68 941 corresponding to the district of *Plumas Unified*. This district has a population of 1 895 households and a median income of 42 770. The value $68 941 is obviously an error of coding. We replaced that observation by the average computed between the previous year and the following year, $4 739. The gross range of that variable is between $3 000 and $11 000 per year, with a median value of $5 000. For total expenditure, the gross range is between $4 500 and $18 000, with a median value of $8 000.

Table 2: Explaining school expenditure per pupil

Variables	Total expenditure per pupil: Rich regime		Poor regime		Instructional expenditure per pupil: Rich regime		Poor regime	
Intercept	4.91	(14.91)	5.74	(19.32)	4.51	(12.07)	4.82	(12.43)
Polarisation	0.025	(0.21)	-0.27	(-3.34)	0.030	(0.23)	-0.24	(-2.70)
Gini	-0.49	(-3.39)	0.001	(0.00)	-0.55	(-3.58)	-0.096	(-0.74)
log(Inc/Hsize)	0.039	(1.46)	0.039	(1.53)	0.074	(2.52)	0.11	(3.21)
Local Tax rev	0.35	(18.39)	0.13	(9.93)	0.32	(16.13)	0.10	(7.13)
Federal Tax rev	0.14	(6.34)	0.28	(14.84)	0.099	(4.04)	0.26	(12.94)
$\tilde{\beta}$	0.189 (0.007)				0.190 (0.007)			
σ^2	0.0247 (0.157)				0.0296 (0.172)			
Size	222		502		223		501	
Mean income	229 615		76 660		229 254		76 426	
Polarisation	0.204		0.338		0.205		0.338	
Gini	0.413		0.407		0.413		0.407	

Figures between parentheses are the estimated Student ratio for the regression parameters and the standard deviations for $\tilde{\beta}$ and σ^2. Student ratio corresponds to the usual star system as follows: *** means 3.29, ** 2.58 and * 1.96.

mean income.[17] However by introducing tax revenue, we hope to reduce the uncertainty contained in the theoretical model for $b_1 < \beta < b_2$, so that a two-regime model should be sufficient. Inference results are provided in Table 2, using a uniform prior information over [0.1,0.4] for the threshold and a diffuse prior on all other parameters.

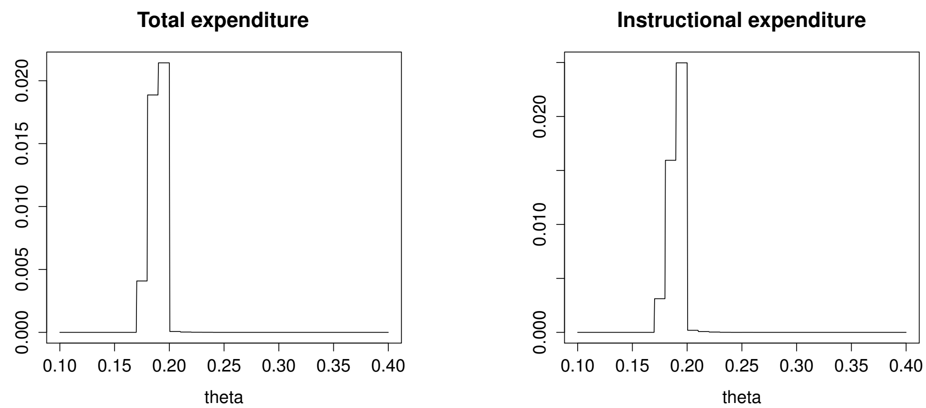


Figure 10: Posterior sample separation
according to the relative proportion of poor households

[17]Income is useful as a conditioning variable for nominal anchoring and also because it is an indirect way of introducing a last characteristic of our income distribution with α. As a matter of fact, we use the mean income calculated from the estimated values of g, β and α.

Importantly whatever the endogenous variable (total versus instructional expenditure), the results are quite similar. The change point between the two regimes always corresponds to a relative proportion of poor households of 0.190 that is very precisely determined. Of course, there are less observations in the rich regime, and on average households are three times richer in the first regime than in the second regime. These values are consistent with the opting-out threshold found above.

This two-regime model says that there are clearly two mechanisms explaining school expenditures. First and quite naturally, federal tax revenues have a greater impact in the "poor" regime than in the "rich" regime, while local tax revenues dominate in the "rich" regime. The impact of household income (normalised by average household size) is significant mainly in the model explaining instructional expenditure and much smaller impact than the impact of tax revenues. We predicted a negative impact of polarisation when there is a high proportion of poor households. This is confirmed from the data, whatever the endogenous variable. The impact of polarisation is positive when the proportion of poor households is low, but this effect is never significant. In this case, inequality, as measured by the Gini index, has a strong negative effect. The result of de la Croix and Doepke (2009) (negative impact of inequality on public schooling expenditure) is valid only for rich districts where the mean household income is around \$230 000. The middle class, as defined here, is clearly not well represented in this regime. Surprisingly the mean *Public expenditures per Pupil* is not very different between the rich regime and the poor regime. We have \$8 609 for total expenditure (\$5 402 for instructional expenditure) in the rich regime and \$8 587 (\$5 132) in the poor regime. So despite the huge difference in mean household income, the effort on public education is quite similar. Only the mechanism determining this quantity is very different. Perhaps what we see here is the compensating effect of federal tax revenues.

When the relative proportion of poor is greater than a threshold, polarisation matters a lot for explaining the level of public spending on education. The average household income in this regime is around \$76 500, which corresponds to the centre of the middle class. Inequality is less marked in this regime (Gini is 0.407) and plays no significant role in explaining public spending. In contrast, polarisation is more marked (0.338) and has a strong negative impact on public spending. So depending on the proportion of poor households, the key variable is either inequality or polarisation and both have a negative impact. Clearly, polarisation and inequality provide complementary information.

In Figure 11, we provide the graph of the mean income distribution in the two regimes, estimated in a non-parametric way, conditionally on the

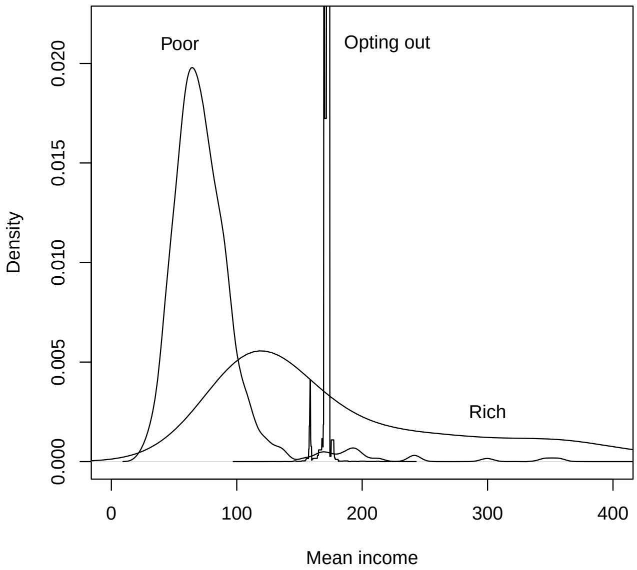


Figure 11: Mean income distribution in the two regimes

Bayesian sample separation. When β is greater than the threshold (poor regime), the income distribution is very concentrated around lower values which cover both the poor and the middle class. There is, however, a very thin long right tail which certainly explains the higher level of income polarisation. When β is lower than the threshold, the income distribution is shifted to the right to cover the upper middle class and the rich households (rich regime). The right tail decays in power slowly and regularly, contrary to the right tail of the poorer regime, which explains the difference in polarisation. The crucial difference between the two regimes lies in the size of the segment around \$172 000, which coincides exactly with our estimated opting-out threshold (opting-out on the graph).

6 Conclusion

In this paper we studied the important relationship between income polarisation, schooling choice and education politics. From a theoretical perspective, we extended the pioneering model developed by de la Croix and Doepke (2009) by allowing for an enriched income distribution which includes in its formulation exogenous parameters describing both income polarisation and income inequality. With respect to the previous literature, our main contribution lies in proposing an innovative mechanism able to capture the consequences of the complex shape of household income distribution on public policies and schooling quality .

The first theoretical result is that having a substantially higher proportion of low-income than of high-income families in a community negatively impacts the quality of public schooling, proxied in the paper by public spending per-pupil. This finding is in line with the empirical evidence which shows that low quality public schools are mainly concentrated in poor areas. Our main contribution lies in analysing the effect of income polarisation, measured as the decline of the middle class, on voted public policies at community level. We show that an increase in income polarisation leads to lower (higher) public education spending per pupil in school districts mainly populated by low (high) income households. Therefore, the effect on public schooling is ambiguous and depends on the particular composition of each school district. This result suggests that income polarisation, as well as income inequality, should be taken into account in the analysis of education politics.

We tested our theoretical conclusions using micro-data from 724 California school districts for 2011-2012. We explain, using the ACS and ELSI databases, that polarisation and inequality are two complementary phenomena and that inequality alone is not fully able to explain public schooling quality. Basically, we used a two-regime switching regression model which proved to be empirically relevant to confirm two features of our theoretical model. A Bayesian approach helped to overcome the difficulties posed by the classical approach to this particular non-linear model, where the usual asymptotic theory does not apply. First we were able to estimate the opting-out threshold a household income threshold above which a household will prefer private schooling. This threshold is fairly high, beyond the income boundaries of the middle class, which is not surprising as California is both a rich state and a state where public schools are of a fairly high quality. Second, we find that the level of public education spending, can be explained by the relative proportion of poor households in a district. For richer districts, we confirm the result of de la Croix and Doepke (2009) that inequality has a negative impact on public spending. However we also find that for poorer districts, polarisation matters in the same way as inequality did for richer districts. These two mechanisms are precisely identified and complemented by the impact of federal tax revenues supplementing local tax revenues in poorer districts. Therefore, depending on the relative proportions of poor and rich households, the key variable is either inequality or polarisation, and both have a negative impact. Polarisation and inequality are key complementary phenomena explaining public education quality.

We validated our theoretical model on Californian data, and this could be seen as a limitation. California is a large and densely populated state. However, it is difficult to speak about the disappearing middle class in California; at most we can speak about the disappearing upper middle class. A more representative state might be Florida, another large and densely populated state, but one with a lower mean household income and a much lower cost of living.

A Proofs

A.1 Proof of lemma 1

Proof *Lemma 1 can be easily proved by deriving (7) with respect to* β*, that is* $\partial\Psi/\partial\beta = g\left(\tilde{x}/x_1 + (x_2/\tilde{x})^\alpha\right) > 0$.

A.2 Proof of lemma 2

Proof *Proceeding as in Lemma 1, we derive equation (7) with respect to* g*. We obtain that* $\frac{\partial\Psi}{\partial g} = \tilde{x}\left(\frac{1}{x_1 - x_2} + \frac{\beta}{x_1}\right) - \left(\frac{x_2}{\tilde{x}}\right)^\alpha (1-\beta)$*. The latter is positive (negative) if and only if* $\beta > \hat{\beta}$ $(\beta < \hat{\beta})$*, with* $\hat{\beta}$ *defined in Lemma 2.*

A.3 Proof of lemma 3

Proof *To prove Lemma 3 it is sufficient to check the first derivatives with respect to* β*. Using (7),* $\frac{\partial\Psi_t}{\partial\beta} = g\frac{\tilde{x}}{x_1} + \left(\frac{x_2}{\tilde{x}}\right)^\alpha > 0$*. Since* $\frac{\partial\tau[\Psi]}{\partial\beta} = \frac{\gamma\eta\frac{\partial\Psi}{\partial\beta}}{1+\gamma\eta\Psi}$*, it follows directly that* $\frac{\partial\tau[\Psi]}{\partial\beta} > 0$*. The derivative* $\frac{\partial\mu}{\partial\beta} = \frac{1}{2}g\left(x_1 - \frac{2x_2\alpha}{\alpha-1}\right)$ *is always negative because* $\alpha > 1$ *and* $x_2 > x_1$ *by assumption. Therefore,* $\frac{\partial s[\Psi,\mu]}{\partial\beta} = \frac{\eta\phi\left((1+\gamma\eta\Psi)\frac{\partial\mu}{\partial\beta} - \gamma\eta\mu\frac{\partial\Psi}{\partial\beta}\right)}{(1+\gamma\eta\Psi)^2} < 0$.

A.4 Proof of proposition 1

Proof *Consider* β *as given. Since* $pol = 4g\beta(1-\beta)$*, we can rewrite* $g \equiv g[pol,\beta]$*,* $\Psi \equiv \Psi[pol,\beta]$ *and* $\mu \equiv \mu[pol,\beta]$*. First of all notice that* $\frac{\partial g}{\partial pol} > 0$ *for all* $\beta \in (0,1)$*. Notice also that* $\frac{\partial\Psi}{\partial pol} \equiv \frac{\partial\Psi}{\partial g}\frac{\partial g}{\partial pol}$ *and* $\frac{\partial\mu}{\partial pol} \equiv \frac{\partial\mu}{\partial g}\frac{\partial g}{\partial pol}$*. The sign of these derivatives depends on the effect of parameter* g *on* Ψ *and* μ*. Deriving (7) with respect to* g*, we get that* $\frac{\partial\Psi_t}{\partial g} > 0$ *if and only if* $\beta > \hat{\beta}$ *as defined in Lemma 2, while* $\frac{\partial\mu}{\partial g} > 0$ *if and only if* $\beta < \bar{\beta} \equiv \frac{x_1(1-\alpha)+x_2(1+\alpha)}{x_1(1-\alpha)+2\alpha x_2}$.

Notice that $\frac{\partial\tau[\Psi]}{\partial pol} \equiv \frac{\partial\tau[\Psi]}{\partial g}\frac{\partial g}{\partial pol} = \frac{\gamma\eta\frac{\partial\Psi}{\partial g}}{1+\gamma\eta\Psi}$*. We then observe that the sign of the derivatives* $\frac{\partial\Psi}{\partial g}$ *and* $\frac{\partial\tau[\Psi]}{\partial g}$ *is the same. The derivative of public schooling expenditure, (16), with respect to income polarisation is given by:* $\frac{\partial s[\Psi,\mu]}{\partial pol} \equiv \frac{\partial s[\Psi,\mu]}{\partial g}\frac{\partial g}{\partial pol} = \frac{\eta\phi\left((1+\gamma\eta\Psi_t)\frac{\partial\mu}{\partial g} - \gamma\eta\mu\frac{\partial\Psi}{\partial g}\right)}{(1+\gamma\eta\Psi_t)^2}$*. The sign of this derivative depends on the sign of the numerator. Of course, it is ambiguous if both derivatives* $\frac{\partial\Psi}{\partial g}$ *and* $\frac{\partial\mu}{\partial g}$ *have the same sign.*

Assume $\bar{\beta} < \hat{\beta}$. When $\beta < \bar{\beta} < \hat{\beta}$ we get $\frac{\partial \mu}{\partial g} > 0$ and $\frac{\partial \Psi}{\partial g} < 0$. Therefore $\frac{\partial s[\Psi,\mu]}{\partial pol} > 0$. When $\beta > \hat{\beta} > \bar{\beta}$, we derive $\frac{\partial \mu}{\partial g} < 0$ and $\frac{\partial \Psi}{\partial g} > 0$. It follows directly that $\frac{\partial s[\Psi,\mu]}{\partial pol} < 0$. However, the sign of the derivative $\frac{\partial s[\Psi,\mu]}{\partial g}$ is ambiguous when $\beta \in [\bar{\beta}, \hat{\beta}]$, because $\frac{\partial \Psi}{\partial g} < 0$ and $\frac{\partial \mu}{\partial g} < 0$.

Assume now $\bar{\beta} > \hat{\beta}$. When $\beta < \hat{\beta} < \bar{\beta}$ we derive $\frac{\partial \mu}{\partial g} > 0$ and $\frac{\partial \psi}{\partial g} < 0$. Therefore $\frac{\partial s[\Psi,\mu]}{\partial pol} > 0$. When $\beta > \bar{\beta} > \hat{\beta}$, we derive $\frac{\partial \mu}{\partial g} < 0$ and $\frac{\partial \psi}{\partial g} > 0$. It follows that $\frac{\partial s[\Psi,\mu]}{\partial pol} < 0$. Again, the sign of the derivative $\frac{\partial s[\Psi,\mu]}{\partial g}$ is ambiguous when $\beta \in [\hat{\beta}, \bar{\beta}]$, since $\frac{\partial \Psi}{\partial g} > 0$ and $\frac{\partial \mu}{\partial g} > 0$.

References

Alesina, A. and Glaeser, E. L. (2004). *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press, Oxford.

Andreski, S. (2006). *Max Weber's Insights and Errors*. Routledge, London.

Arcalean, C. and Schiopu, I. (2016). Inequality, opting-out and public education funding. *Social Choice and Welfare*, 46(4):811–837.

Bauwens, L., Lubrano, M., and Richard, J.-F. (1999). *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, Oxford.

Booza, J. C., Cutsinger, J., and Galster, G. (2006). Where did they go? The decline of middle-income neighborhoods in metropolitan America. *The Brooking Institution*, pages 1–23. Metropolitan Policy Program.

Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100(1):1–40.

Corcoran, S. and Evans, W. N. (2010). Income inequality, the median voter, and the support for public education. Working Paper 16097, National Bureau of Economic Research.

de la Croix, D. and Doepke, M. (2009). To segregate or to integrate: education politics and democracy. *The Review of Economic Studies*, 76(2):597–628.

Easterly, W. (2001). The middle class consensus and economic development. *Journal of Economic Growth*, 6(4):317–335.

Esteban, J.-M. and Ray, D. (1994). On the measurement of polarization. *Econometrica*, 62(4):819–851.

Foster, J. E. and Wolfson, M. C. (2010). Polarization and the decline of the middle class: Canada and the U.S. *The Journal of Economic Inequality*, 8(2):247–273.

Gastwirth, J. L. (1972). The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, 54(3):306–316.

Hansen, B. (2000). Sample splitting and threshold estimation. *Econometrica*, 68(3):575–603.

Levy, F. and Murnane, R. J. (1992). U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations. *Journal of Economic Literature*, 30(3):1333–1381.

Melindi-Ghidi, P. (2016). Inequality, educational choice and public school quality in income mixing communities. Working Paper 29-2016, Aix-Marseille School of Economics.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927.

Mills, C. W. (1951). *White Collar: The American Middle Classes.* Oxford University Press, Oxford.

Noss, A. (2013). Household income: 2012. American Community Survey Briefs ACSBR/12-02, U.S. Census Bureau.

Persson, T. and Tabellini, G. E. (2002). *Political Economics: Explaining Economic Policy.* MIT Press, Cambridge MA.

Piketty, T. and Saez, E. (2003). Income inequality in the United States: 1913-1998. *Quarterly Journal of Economics*, 118(1):1–39.

Quandt, R. E. (1966). Old and new methods of estimation and the Pareto distribution. *Metrika*, 10(1):55–82.

Renwick, T. and Short, K. (2014). Examining the middle class in the United States using the lens of the supplemental poverty measure. Working Paper 2014-30, U.S. Census Bureau.

Schader, M. and Schmid, F. (1994). Fitting parametric Lorenz curves to grouped income distribution: a critical note. *Empirical Economics*, 19(3):361–370.

Soares, J. (2003). Self-interest and public funding of education. *Journal of Public Economics*, 87(3-4):703–727.

von Hippel, P. T., Scarpino, S. V., and Holas, I. (2015). Robust estimation of inequality from binned incomes. *Sociological Methodology*, 46(1):212–251.

Wolfson, M. C. (1994). When inequalities diverge. *The American Economic Review*, 84(2):353–358.