Abstract

Do communities with the same level of inequality but 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. We show that not only income inequality but also income polarisation matters in explaining differentials in public school spending across communities. Indeed, public schooling is an important concern of the middle class which is more inclined to pay higher level of taxation in order to have better public schools. On the contrary, poorest households might put less interest in public education while rich parents are more willing to opt out from 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.

Acknowledgments: This work has been 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 an increase in the importance of the middle class (see e.g. Levy and Murnane 1992). Essentially, the middle class was constituted by educated white collars, ranging from permanent teachers to journalists and lawyers.¹ This upward social mobility was made possible thanks to the importance of schooling and college education. The quality of schooling and consequently (local) public spending became of a very important concern for the middle class, because it has seen its immediate social consequences for itself (Alesina and Glaeser 2004, Chap 7). This importance 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 income of the upper deciles increased much rapidly compared to the lower deciles. This movement was eventually accompanied by a lesser social mobility.

Booza et al. (2006) underline 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 importance 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 polarisation, the middle class being obliged to join poor neighbourhoods while the richer part of the population was concentrated in very rich districts.

These changes in income distribution and in social mobility can be analysed using the data of the PSID between 2009 and 2011 by estimating a quantile income transition matrix. In Table 1, we divide household income (normalised by the squared root of household size) into quantile intervals. The first interval (< 20%) corresponds grossly to the poor households. We detail the 20% upper part of the income distribution in three groups. The last class (> 95%) corresponds to the rich, while the previous interval (90%-95%) represents the upper middle class. We find as usual that the poor and the rich are rather immobile because they have a high probability of staying where they are. But what is remarkable in this transition matrix is that the upper middle class has much greater chances of falling down than of climbing up. Mobility is minimum not far from the median (60%-80%), with slightly more chances of falling down than getting up. But after climbing up, that

¹See for instance Mills (1951) for a tentative definition of the middle class.
part of the middle class (80%-95%) has much more chances to fall down than to climb up.

Table 1: Income transition matrix for the USA

<table>
<thead>
<tr>
<th>PSID 2009-2011</th>
<th>&lt; 0.20</th>
<th>0.2-0.4</th>
<th>0.4-0.6</th>
<th>0.6-0.8</th>
<th>0.8-0.9</th>
<th>0.9-0.95</th>
<th>&gt; 0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.20</td>
<td>0.65</td>
<td>0.23</td>
<td>0.09</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.2-0.4</td>
<td>0.23</td>
<td>0.44</td>
<td>0.23</td>
<td>0.08</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>0.4-0.6</td>
<td>0.08</td>
<td>0.22</td>
<td>0.45</td>
<td>0.20</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>0.6-0.8</td>
<td>0.03</td>
<td>0.07</td>
<td>0.19</td>
<td>0.53</td>
<td>0.14</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>0.8-0.9</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.27</td>
<td>0.44</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>0.9-0.95</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.07</td>
<td>0.26</td>
<td>0.41</td>
<td>0.18</td>
</tr>
<tr>
<td>&gt; 0.95</td>
<td>0.01</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.09</td>
<td>0.13</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Maurin (2009), working on French data around the same period, found that even if social mobility was unchanged in France after 1993, the middle class experienced a real fear of falling down. In this mobility context, inequality should no longer be perceived as an opportunity, but as a risk. This pattern in social mobility should have a strong influence on the willingness to pay taxes for local public spending and finally schooling quality.

Several papers have stressed the relation between income inequality and public school quality. First of all, there has been a traditional incentive to have public schools of good quality. Card and Krueger (1992), using US data over 1920-1949 illustrate the influence of school quality on the rate of return to 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 conclude to 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 two systems of education compete, private schools and public schools. The proportion of private enrolment varies greatly over countries as well as the type 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 going to be affected. de la Croix and Doepke (2009), on the contrary, propose a model (and an empirical test) where income inequality is the main determinant for segregation in the schooling system. This will not affect public funding.
for public school if all classes have a proportional representation (which is coherent with the median voter theory). However, if the richer class manages to dominate ideologically, the political system will be biased to the rich and a probabilistic voting model can lead to an equilibrium where the two ends meet against the middle class to vote against public funding as in Epple and Romano (1996).\footnote{Alesina and Glaeser (2004, Chap 7) have shown that ideology plays a very important role in the shaping of public schools over the two sides of the Atlantic.} Along these lines, Melindi-Ghidi (2016) shows that these results persist even though the model allows for the geographical mobility of households.

The idea we would like to illustrate 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 the income distribution cannot be summarised by this single indicator. If income polarisation was for long recognised as the disappearance of the middle class by many sociologists after Max Weber (Andreski 2013, p. 105), the proof that inequality was a different concept from polarisation was first given in Wolfson (1994) while Esteban and Ray (1994) provided an axiomatisation of it. Indeed, income inequality characterises the dispersion of an income distribution. It can be measured by various indices such as for instance the coefficient of variation or the Gini index. Differently, income polarisation characterises the increase of the ends of a distribution at the expense of its centre and both its definition and 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 can be the source of potential conflicts (Esteban and Ray 1994) and leads to an uneven society while a middle class consensus would lead to more economic development and more investment in human capital as advocated in Easterly (2001).

From a theoretical perspective, the model developed in this paper is an extension of the pioneeristic setting developed by de la Croix and Doepke (2009). Compared to their work, we assume a more realistic income distribution. This important extension allow us to include in their model a measure of income polarisation within school districts. While the original model is able to capture the effect of variations of the income dispersion on public education policies, it is not able to explain if these policies are also affected by changes in income distribution that do not alter the dispersion of income within districts. Put differently, if income distribution is represented with a standard uniform distribution, the role of income polarisation on public education policies and school choices cannot be taken into account. Since the contribution of this paper is to study the main consequences of income
polarisation on public education policies, we introduce a parametric form for an income distribution which could include in its formulation parameters explaining this phenomenon. Thus, we propose a hybrid mixture model formed by two uniform and a Pareto distributions.\(^3\)

In a recent paper, Arcalean and Schiopu (2016) assume that income is distributed accordingly to a Pareto distribution. This extension was important to capture the empirically relevant increase in income inequality we observe in the U.S. data. The authors are able to obtain a flexible parameterisation of the income distribution showing that the effects of higher income dispersion on public education spending can be non-monotonic in the mean income of the economy. However, as for de la Croix and Doepke (2009), this model cannot explain the important role of income polarisation on public policies and schooling choices. Indeed, differing school districts can be characterised by the same level of income dispersion, but very different internal population composition in terms of income classes. Of course, the presence of a majority of low- or high-income households has an important impact on school choice, public and private enrolment rates and political behaviour.

The model developed in this paper is able to keep into account the dispersion of the income distribution as well as its complex shape. 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 an important phenomenon: *income polarisation*. Indeed, the main theoretical contribution of the paper is to analyse the consequences of differences in the income composition of school districts on public education policies and school choice. We show that the effect of a shrink of the middle class relative to the ends is ambiguous and depends on the population composition of each school district. In particular, in districts largely populated by poor compared to rich households, income polarisation is more likely to negatively impacts public schooling quality. The reason comes from the fact that a positive variation in participation rate in public education as a consequence of increasing polarisation is not compensated by the positive effect on taxation to finance the public education system. By contrast, in school district populated by a large share of rich compared to poor families, income polarisation pushes down participation rate to public school because rich parents are more willing to enrol their children in private schools since education is a normal good. Even though taxation rate goes down, the participation effect might dominate the taxation effect in absolute terms and, therefore, income polarisation

\(^3\)To this end, we will need an income distribution composed of three income classes: the poor, the middle class and the rich.
will positively impact the quality of public schooling.

Using the ACS and ELSI databases, we aim to confirm this hypothesis through a two regime regression model. More precisely we want to test if polarisation plays or not a role on public spending and thus on public school quality. We find that in school districts largely populated by poor households income polarisation matters, while in the opposite situation, where it is the rich households that dominate, inequality plays the most important role in explaining public school spending.

The paper is organised as follows: section 2 motivates our objectives with important stylised facts for the State of California. In this section we also introduce an enriched income distribution that allows to take into account income polarisation in the analysis. In section 3 we extend the theoretical model of de la Croix and Doepke (2009) with the proposed income distribution. In this section we study the theoretical impact of income polarisation on public education policies and school choice. Section 4 focuses on the estimation of our enriched mixture distribution and indices. Section 5 concentrates on an empirical analysis of the main theoretical results. Section 6 concludes.

2 From Inequality to Polarisation

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 distribution. In this section we want to precise the relationship between school quality and the shape of the income distribution with a clear distinction between inequality and polarisation. Using California school district data for year 2011-2012, we analyse the role played by income polarisation compared to inequality. We first propose some stylised facts, on public school quality and its relation to income. Second, we present a model for an income distribution which is more general than the uniform distribution considered in de la Croix and Doepke (2009), but allows both for analytical results and an empirical measure of income polarisation.

2.1 The ACS and ELSI Databases

The American Community Survey (ACS)\textsuperscript{4} 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

\textsuperscript{4}Available on: http://nces.ed.gov/programs/edge/demographicACS.aspx
(ELSI)\(^5\) which provides information on public and private schools.\(^6\) We restrict our analysis to households with enrolled children in schools (public or private). After merging these two data bases, our sample concerns 724 school districts, documenting information on income distribution over the school district (grouped data for fixed classes of income), education (pupil/teacher ratio, instructional expenditure, private enrolment...), local taxes, fertility, and many other items. We now analyse some stylised facts concerning school choice, the influence of inequality versus polarisation on school quality. In the final section of this paper, we use this data set to test empirically the theoretical model of section 3.

2.2 Stylised Facts

When it comes to schooling, parents often have to work out whether to send their children to private or public schools. Different factors impact their decisions. In the US, public schools are financed through Federal, State, and Local taxes while private schools funding comes from a variety of sources including tuition fees which can be very high. Public schools are offered free of charge. This important difference might create a gap between the wealthiest and poorest households since only the wealthier have the financial means to provide private education to their children. Therefore, this system can generate school segregation not only between but also within school districts.

More precisely, public schools are managed at the level of school districts which are governed by elected councils. It is a form of local government placed under the responsibility of the State (California in our case study). The elected council has the authority for managing local public schools. Considering Californian school districts, we observe a clear income-related segregation at school district level. Figure 1 shows the empirical relation between the share of private enrolment and households 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 enrolment (nearly zero \(p\)-values for the Pearson correlation coefficient). This very high correlation may be a first explanation in the varying share of children sent to private school for each school district.


\(^6\)These two data bases are particularly interesting as they provide information at the level of school districts 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 restricted groups of children and parents.
Examining the data, the minimum share of private enrolment over Californian school districts is 0.006, which is very low, while the highest one is 0.727, which is very high. The median rate is comparatively low with 0.116. We effectively observe school segregation in some districts, but this is not a general phenomenon. There is at least within each school district one household that choose private school for its children and at least one that choose public school. Presumably because school quality is rather high in California, we do not observe in our database a school district where only one alternative is chosen by all households.

Parents who can afford private schools may prefer to opt-out from the public school system because they expect a higher schooling quality. The latter can be measured in different ways. The number of pupils per class has long been used to identify school quality, starting with the Coleman report (1966). This is measured by the student-to-teacher ratio. More recent literature, such as de la Croix and Doepke (2009), prefer to use expenditure variables: school quality is measured by the level of instructional or total expenditures allocated to public schools. What makes that some school districts benefit from more financial means than others? As explained before, partly because of the level of local taxes. That level is voted by the parents living in the school district and thus depends on the income distribution of that district.
If we find in the literature many references concerning the relation between income inequality and school quality, the direction of the link is not clear, especially when public and private schools are competing (see the references given in the introduction). Figure 2 provides two plots to illustrate the relation between public school quality and inequality measured by a Gini coefficient. On the left side we plot the empirical relation between public instructional-expenditure and the Gini coefficient while on the right side we plot the empirical relation between the public pupil-teacher-ratio and the Gini.\textsuperscript{7} What we observe at the level of our school district data for California is a total absence of relation between inequality and school quality. We get very low and non significant correlation coefficients with \(p\)-values equal respectively to 0.81 and 0.61.

Figure 2: Relation between public school quality and Gini coefficient.

However, focusing only on inequality may lead to a misunderstanding of the rich 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 a totally different importance of the middle class. This figure illustrates the difference between polarisation and inequality. We shall propose below a model for a slightly

\textsuperscript{7}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”.

9
more general modelling of the income distribution which allows both analytical results and the possibility of income polarisation. Polarisation brings in fact a lot of new informations. Let us consider again Figure 2, but replacing the Gini index by a measure of polarisation that we shall develop below. We obtain Figure 4. Now the correlation between public school spending and our new indicator becomes strongly significant. More in details, in the left side of Figure 4, relating polarisation and instructional expenditure, the Pearson’s product-moment correlation is equal to -0.19 with a near zero p-value. In the right side of Figure 4 which concerns polarisation and the pupil-teacher ratio, the Pearson’s product-moment correlation is equal to 0.16 with also a near zero p-value. We have thus a clear negative influence of polarisation.
on school quality when measured by both indicators. The more polarised a school district is, the highest the proportion of poor and rich and relatively to the middle class. The former being the more inclined to favour high public school quality, it explains the negative relationship. This motivates the development of a new income distribution that allows both for a simple measure of polarisation and of analytical results for our theoretical model developed in section 3.

2.3 An Enriched Income Distribution with Polarisation

In this section, we concentrate on a parametric form for an income distribution which could include in its formulation parameters monitoring directly 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. For this, 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 over $x_2$. Income is supposed to follow a uniform distribution inside each class. We have thus a mixture of three uniform distributions. If $g$ is a parameter monitoring the importance of the middle class and $\beta$ a parameter monitoring the relative balance between the poor and the rich, we have:

$$f(x) = \frac{g\beta}{x_1} 1(x < x_1) + \frac{1 - g}{x_2 - x_1} 1(x_1 \leq x < x_2) + \frac{g(1 - \beta)}{x_{max} - x_2} 1(x_2 \leq x < x_{max}),$$

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

$$f(x) = \frac{g\beta}{x_1} 1(x < x_1) + \frac{1 - g}{x_2 - x_1} 1(x_1 \leq x < x_2) + \frac{g(1 - \beta)}{x_{max} - x_2} 1(x_2 \leq x < x_{max}) + g(1 - \beta) \frac{\alpha x_2^\alpha}{x^\alpha + 1} 1(x \geq x_2).$$
Equipped with these two distributions, we can now explain the relation between income polarisation and inequality.

2.4 Polarisation and Inequality in our Mixture Income Distribution

Let us now discuss polarisation in the context of our mixture model. The shape of the centre of the distribution, and consequently the importance of the middle class, is monitored by the value of $g \in [0,1]$. Its relative importance 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 $\beta$. For $\beta = 0$ we have no poor, and for $\beta = 1$ we have no rich. For these two extreme cases, we cannot have polarisation, because one of the extreme group disappears. For $\beta = 0.5$, the distance between the poor and the rich is maximum. Combining these two aspects, we propose a polarisation index at value in $[0,1]$ 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 importance of rich and poor). It is 0 when either $g = 0$ or when either one group rich or poor disappears ($\beta = 0$ or $\beta = 1$).

Polarisation and inequality are nevertheless linked, while being however 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

Figure 5: Relation between polarisation and inequality in California
is equal to 0.29 (and a nearly zero p-value) for this sample. However, the slope of regression line is far from 1. And the plot is rather scattered. Which means that there are districts having the same level of inequality and totaly 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 build by de la Croix and Doepke (2009). The authors develop a model of school choice and voting on public education, assuming that income is distributed accordingly 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 \(\sigma\), on the quality of the public school system. However, this simplifying assumption on income distribution does not allow to describe the distribution we observe in reality within U.S. school districts as well as to explain why in school districts with the same level of inequality we can observe differentials in public school quality. For these reasons, we assume that the distribution of income is characterised by a mixture of two uniform and a Pareto distribution.\(^8\) Indeed, the assumption of a uniform income distribution cannot account for the role that income polarisation might have in explaining differentials in schooling quality. Our objective is to understand the main effects that an income distribution that accounts for differences between and within social classes might have on public policies, and therefore, on schooling quality and segregation.

3.1 Households’ Problem: the Theoretical Set-up 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 either in public or private schools. The parameter \(\gamma \in \mathbb{R}^+\) is the overall weight attached to children and \(\eta \in (0, 1)\) is the relative weight of quality. Public education is free of charges, while private education can be acquired by paying a tuition

\(^8\)In section 4, we estimate the California income distribution showing that this mixture income distribution is a good representation of the analysed date. See Figure 6.
fee. When fertility and education decisions are taken, households vote for a level of income tax rate to finance public education spending.

A representative agent has an additive and separable utility function:

\[ u = \ln(c) + \gamma[\ln(n) + \eta \ln(h)], \]  

with \( h = \max\{s, e\} \) representing the level of child human capital, that is, the quality of education acquired by each child. The variable \( s \) represents the quality of public school while \( e \) represents the private investment on education. Public and private education are mutually exclusive. Since private education spending is assumed tax deductible, the budget constraint writes:

\[ c = (1 - \tau)[x(1 - \phi n) - n e], \]

where \( \tau \) is the income tax rate, \( x \) the exogenous income level and \( \phi \) the fixed raising cost of a child. 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 a given choice of education type:

\[
\text{Public: } \quad e^s = 0, \quad n^s = \frac{\gamma}{\phi(1 + \gamma)} \\
\text{Private: } \quad e^e = \frac{x\eta\phi}{1 - \eta}, \quad n^e = \frac{\gamma(1 - \eta)}{\phi(1 + \gamma)}.
\]

Observe that in general \( n^s > n^e \): parents choosing the public system have more children. Of course, fertility rates do not depend on income, due to the logarithmic form of the utility function. We define the indirect utility function \( V^s \), corresponding to the choice of public education, and \( V^e \), corresponding to the choice of private education, by replacing the budget constraint (4) and the optimal decisions (5) in the utility function (3). The final education 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 can happen only if the agent has an income greater than a threshold given by:

\[ x > \bar{x} \equiv \frac{E[s](1 - \eta)^{\frac{\eta - 1}{\tau}}}{\phi\eta}, \]

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

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 households’ problem developed by de la Croix and Doepke (2009) to help the reader to follow the theoretical extension proposed in our paper.

Since \( e^s = 0 \), for simplicity we will define \( e^e = e \) in the paper. A regression of the public enrolment rate on the average household size has a t test of 5.42.
that is Lemma 2 in de la Croix and Doepke (2009).

Given an income level $x$, the larger the expected quality of public schooling, $E[s]$, the lower the probability to opt-out from the public school system. Since households have perfect foresight over the outcome of the political process and, consequently, over the policies adopted by the government, $E[s] = s$.

### 3.2 Introducing Income Polarisation

Let us now consider our proposed income distribution (2) which is a mixture of two uniform and a Pareto distributions. We first determine the participation rate in the public school system $\Psi$ as the integral between 0 and $\tilde{x}$ of the income distribution:

$$\Psi = \int_{0}^{\tilde{x}} f(x)dx = g\beta \frac{\tilde{x}}{x_1} + (1-g)\frac{\tilde{x}}{x_2} - (1-\beta)g \left(\frac{\tilde{x}}{x_2}\right)^{-\alpha} \quad (7)$$

As in de la Croix and Doepke (2009), fertility and education choices are determined before the political process takes place. It follows that the opting-out threshold $\tilde{x}$ is taken as given. Therefore, the derivative of (7) with respect to $\beta$ under the perfect foresight assumption is always positive.

**Lemma 1** An increase in the relative proportion of poor, $\beta$, compared to rich households positively affects the participation rate in public school.

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

However, the effect of parameter $g$ on the participation rate is 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** An increase in the share of the ends compared to the middle class, has an ambiguous effect on the participation rate in public school:

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

with

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

Lemma 2 indicates that if the presence 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 the reason comes from the observation that education is a normal good because richer households are more demanding in terms of it. Therefore, poor households are more reluctant to opt-out from the public school system and invest in private education for their children, as shown in Lemma 1. In other words, the internal composition of each school district matters, even though the dispersion of the income at district level is the same.

3.3 Equilibrium and Income Polarisation under Probabilistic Voting

In this section we 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 at value between $1 - \sigma$ and $1 + \sigma$ to the mixture (2) defined over the support $[0, +\infty]$. This generalisation allows us to introduce a measure of income polarisation while the original model can only account for income inequality. The analysis of the effects of the two parameters, $\beta$ and $g$ on the political mechanism and, therefore, on the equilibrium is the main theoretical contribution of this section.

We first observe that as each school district must have a balanced budget, the total spending for public schools given by

$$\int_0^{\tilde{x}} sn^s f(x) dx,$$

(8)

equals the total local income taxation. As taxation bears on both types of households, those sending their children to public school and those sending their children to private school, we get

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

(9)

Since education spending is assumed tax deductible and fertility is endogenous, the taxable income is the same for parents choosing public or private
school and does not depend on the school choice of households. Indeed, using (5), it is easy to verify that \( x(1 - \phi n^s) = x(1 - \phi n^e) - \epsilon n^e \equiv x/(1 + \gamma) \). We can rewrite the balanced budget rule of the local government as follow:

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

We follow de la Croix and Doepke (2009) assuming that voting decisions are taken when fertility and education choices are determined. Therefore, the opting-out threshold \( \tilde{x} \) is taken as given when households vote for public policies.

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, \( \tau \), the participation rate, \( \Psi \), and the mean of the income distribution, \( \mu \):

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

with

\[
\mu \equiv \int_0^\infty xf(x)dx = \frac{1}{2} \left[ (1 - g)(x_1 + x_2) + \frac{2(1 + \beta)gx_2\alpha}{\alpha - 1} + g\beta x_1 \right].
\]

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:

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

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 follow:

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

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

\[
\Omega[\tau] = \ln \left[ \frac{1 - \tau}{1 + \gamma} \right] + \gamma \ln \left[ \frac{\gamma}{\phi(1 + \gamma)} \right] + \gamma\eta \ln \left[ \frac{\mu\tau\phi}{\Psi\gamma} \right] \int_0^{\tilde{x}} f(x)dx
+ \int_0^\infty \ln[x]f(x)dx + \int_\tilde{x}^\infty \left( \gamma \ln[1 - \eta] + \gamma\eta \ln \left[ \frac{x\eta\phi}{1 - \eta} \right] \right) f(x)dx. \tag{14}
\]
Now, taking the first-order condition with respect to $\tau$ for a maximum, allow us to express the voted tax rate in terms of participation rate in the public school system:

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

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

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

Since $\Psi$ and $\mu$ are both functions of parameters $\beta$ and $g$, differently from the existing theoretical literature, our model is able to capture the role of income polarisation on public policies and schooling choice. More precisely, the population composition of each school district might have a crucial impact on public policies also in school districts where the income dispersion, and therefore the income inequality, is at the same level. Put differently, our model is able to explain the empirical case in which differentials in public schooling quality emerge even in school districts where inequality is at similar level.

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

It should be noted that compared 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 school and the mean of the income distribution. In particular, the higher average income of the economy the higher the public schooling spending $s[\Psi, \mu]$. This result is not surprisingly since it is reasonable to expect that richer school districts 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 $\phi, \gamma, \eta$, two school districts with the same level of inequality might perform differently in terms of schooling quality.

For this important theoretical and empirical issue, it is quite crucial to underline the main differences of our setting with respect to 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 public school 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 guarantee 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 might be different in economies with the same level of inequality.

A second crucial difference is that in our set-up both the fully private regime and the fully public regime cannot be an equilibrium outcome. Assuming a more realistic income distribution function over the support \([0, \infty]\) necessarily implies a certain level of schooling segregation. 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 enrolment rate. This means that households always segregate by schooling because they do have the choice between public and private education.

Third, our model is able to explain schooling differentials through the channel of the income polarisation. This is in line with our view for which not only income inequality but also how the income is distributed matters. Indeed, if we consider as given the share of the middle income group, the relative balance between poor and rich households can have important effects on public policies.

**Lemma 3** An increase in parameter \(\beta\) leads to a higher tax rate and a lower public school spending:

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

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

Lemma 3 establishes that given the size of the middle income group, a school district with a larger share of poor compared to rich households have a higher participation rate in public school, a higher tax rate and a lower schooling spending. This theoretical result confirms the empirical observation for which poor urban areas are characterised by public schools of low quality but 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 $\beta$, the effect of an increase in income polarisation on public school spending is negative (positive) if $\beta > \max\{\beta, \bar{\beta}\}$ ($\beta < \min\{\beta, \bar{\beta}\}$) and ambiguous if $\beta \in [\min\{\beta, \bar{\beta}\}, \max\{\beta, \bar{\beta}\}]$.

Proof: Consider $\beta$ as given. Since $pol = 4g\beta(1 - \beta)$, we can rewrite $g \equiv g[\text{pol}, \beta]$, $\Psi \equiv \Psi[\text{pol}, \beta]$ and $\mu \equiv \mu[\text{pol}, \beta]$. First of all notice that $\frac{\partial g}{\partial \mu} > 0$ for all $\beta \in (0, 1)$. Notice also that $\frac{\partial \Psi}{\partial \mu} \equiv \frac{\partial \Psi}{\partial g} \frac{\partial g}{\partial \mu}$ and $\frac{\partial \mu}{\partial g} \equiv \frac{\partial \mu}{\partial \mu}$. The sign of these derivatives depend on the effect of parameter $g$ on $\Psi$ and $\mu$. Deriving (7) with respect to $g$, we get that $\frac{\partial \Psi}{\partial g} > 0$ if and only if $\beta > \bar{\beta}$ as defined in Lemma 2, while $\frac{\partial g}{\partial \mu} > 0$ if and only if $\beta < \bar{\beta} \equiv \frac{\bar{x}_1(1-\alpha)+\bar{x}_2(1+\alpha)}{\bar{x}_1(1-\alpha)+2x_2}$.

Notice that $\frac{\partial \tau[\Psi]}{\partial g} \equiv \frac{\partial \tau[\Psi]}{\partial g} \frac{\partial g}{\partial \mu}$. 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 the public schooling spending, (16), with respect to income polarisation is given by: $\frac{\partial s[\Psi, \mu]}{\partial \text{pol}} \equiv \frac{\partial s[\Psi, \mu]}{\partial g} \frac{\partial g}{\partial \mu} = \frac{\eta \alpha (1+\gamma \Psi) \frac{\partial s[\Psi, \mu]}{\partial \mu} - \gamma \mu \frac{\partial s[\Psi, \mu]}{\partial \Psi}}{\eta \alpha (1+\gamma \Psi)^2}$. 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 $\frac{\partial \tau[\Psi]}{\partial g}$ have the same sign.

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

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

A first important hint of Proposition 1 is the following: 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. Indeed, the effect of a variation of 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 is composed by a sufficiently large proportion of poor compared to rich households, $\beta > \max\{\beta, \bar{\beta}\}$, the relative reduction of the middle income group will negatively impact the public school spending per child. Indeed, in this particular scenario, an increase in the share of the ‘ends’ compared to the middle will positively impact the participation rate in public school. Even though the tax rate increases, our model predicts that in school districts
populated by a large share of poor relative to rich households, the effect of income polarisation on public school spending will be negative.

When school districts are populated by a large share of rich households compared to poor households, $\beta < \min\{\hat{\beta}, \bar{\beta}\}$, participation rate goes down as a consequence of an increase in income polarisation. Indeed, rich parents are more willing to invest in education and enrol their children in private schools. In this scenario, the reduction in the size of the middle income group negatively impacts the participation rate in public school and the voted tax rate. Therefore, as more households opt-out from the public schooling system, the public spending per student increases when the relative balance between rich and poor households is biased towards the right of the income distribution.

However, the effect of income polarisation on the public school spending is found to be ambiguous when the share of poor and rich households is relatively balanced, that is when $\beta \in [\bar{\beta}, \hat{\beta}]$. 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 of the participation rate in public school compared to the variation of the tax rate as a consequence of income polarisation.

**Corollary 1** The poorer the school district, the more likely income polarisation negatively impacts public school spending.

To demonstrate Corollary 1, consider the following. Given $\beta$, we can 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 $k = \frac{\gamma \mu}{1 + \gamma \psi}$. The larger $\mu$, the larger the value of $k$. This implies that for a given participation rate in the public school system, the effect of income polarisation on public school spending is more likely to be positive in rich school districts characterised by high mean income. Moreover, since high income parents are more willing to enrol their children in private schools, the participation rate to public school will be lower in richer school districts. Indeed, the possibility to opt-out from the public education system, the fact that education is a normal good, and the political implication for which poor and rich households vote against redistribution, are behind our result for which income polarisation has a negative (positive) impact on the quality of public school in poor (rich) school districts. In the next two sections, we will check if micro-data for California are able to confirm our theoretical conclusions.
4 Estimation of the California Income Distribution

In the ACS data base, information of the household income distribution (with enrolled children) is provided at the school district level in the form of grouped data represented by ten unequal classes, with top coding for the largest. The lowest class represents the number of households with an income plus benefits below 10 thousand dollars per year, while the largest class corresponds to the number of households with a year income and benefit greater than 200 thousand dollars. It is not possible to apply an equivalence scale for these income data, because we have grouped observations for both income and family composition.\footnote{The matter should not be too crucial because we have restricted our attention to households with children and the average household size does not vary too much over the 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, using the new OECD equivalence scale.} We have represented this distribution for the whole state of California, regrouping all school districts, in Figure 6. We have represented the last open class by a Pareto distribution. The way the Pareto parameter ($\hat{\alpha} = 2.28$) is estimated is explained below in the text.
4.1 A Stylised Income Distribution

We have decided to regroup the ten classes in three groups so as to figure out the poor, the middle and the rich classes and get an empirical model in adequation with our theoretical model. We have chosen to represent the poor households by those who have a year income lower than $25 000. In 2012, the official median household income was $58 328 for California 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) would represent 55\% of that median income. We must note however that there is no relative poverty line in the US. Using $25 000 as a poverty line, Table 2 indicates a percentage of poor of 16\% while the Bureau of Census indicates a rate of poverty for California of 17\% in 2012. So, we are pretty safe. We have now 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 last decile of an income distribution represents both the upper middle class and the very rich. However, among the variety of definitions of the middle class reported in Renwick and Short (2014), those refereing to quantiles use the range between 0.25 and 0.75 of the income distribution. There are also definitions in terms of the 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 limit of $95 216. We have decided to 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. There are 32\% of the households with an income greater than $100 000.

We have represented this stylised income distribution in Figure 7. It is interesting to compare the income distribution of each school district with that reference distribution, remembering that the total number of rich is twice the number of poor. There are 331 districts out of 724 where the middle class

\(^{13}\)This number 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

is dominant (more than 50%) and 125 districts where the rich are dominant (more than 50%). The 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 Parameters 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 last 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 $j = 1, \ldots, 3$ and $n = \sum_{j=1}^{3} n_j$ the total number of households in a school district. For each school district, we estimate first:

\begin{align}
\hat{g} & = 1 - n_2/n, \\
\hat{\beta} & = n_1/(n\hat{g}),
\end{align}

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 value of the 724 \( \beta \) (the relative proportion of poor) is 0.34 with a standard deviation of 0.26. The relative proportion between rich and poor households varies quite a lot.

Let us now turn to the estimation of \( \alpha \) for the Pareto member. Let us define \( nc_i \) as the number of households in each of the ten income classes. We suppose that \( nc_{10} \) represents the number of household with an income greater than \( x_{10} = \$200,000 \). Quandt (1966) explains how to estimate the Pareto parameter in the case of grouped data (see also von Hippel et al., 2015):

\[
\hat{\alpha} = \frac{\log(nc_8 + nc_9 + nc_{10}) - \log(nc_{10})}{\log(x_{10}) - \log(x_8)},
\]

(19)

where \( nc_8, nc_9 \) and \( nc_{10} \) represents 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) because 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 \( nc_9 \). The predetermined class boundaries cannot represent adequately the right tail of the income distribution for these 57 school districts. The last class has to be divided in two income classes: incomes between \$200,000 and \$300,000 for the tenth class and incomes greater than \$300,000 for an 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 \( \tilde{nc}_{10} = nc_{10}/1.4 \). So that the last open class contains \( \tilde{nc}_{11} = (1 - 1/1.4)nc_{10} \). With this extension, we estimate \( \alpha \) as:

\[
\hat{\alpha} = \frac{\log(nc_8 + nc_9 + \tilde{nc}_{10} + \tilde{nc}_{11}) - \log(\tilde{nc}_{11})}{\log(x_{11}) - \log(x_8)}.
\]

(20)

We could estimate \( \alpha \) in 611 cases out of 724\(^4\). The median value is 2.78 with a standard deviation of 1.30. The minimum value is 1.20 and the max 8.00. So inequality among the rich is in general rather low with a median Gini of 0.22, but it can go to the rather high value of 0.40 in 9% of cases.

4.3 An Improved Method using the Mean Income

The mean income of each district is provided in the data set. We can use this information in order to get an improved estimator for \( \alpha \). The mean of

\(^4\)The remaining cases correspond to the situation where \( nc_{10} = 0 \) and is represented by a mixture of three uniforms.
our overall distribution is the weighted sum 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 \( \alpha > 1 \). This formula is an application of the properties of the uniform and Pareto distributions. Our previous estimate of \( \alpha \) together with \( \hat{g} \) and \( \beta \) provides an estimation for the empirical mean which has to be confronted to the mean reported in the data for each school district. We can improve our first estimator of \( \alpha \) by minimising in \( \alpha \) the following loss function

\[ (\mu(\alpha, \hat{g}, \hat{\beta}) - m_s)^2 + (\alpha - \hat{\alpha})^2, \]

where \( \hat{\alpha} \) is the initial estimator resulting from the application of (19) or (20) and \( m_s \) is the empirical mean given in the data set. With the simple method, we have 12 districts where \( \alpha \leq 1 \). Our method is a way to include supplementary information in order to constrain \( \alpha \) to be greater than 1.

### 4.4 Gini Coefficient for Grouped Data

The Gini coefficient measures the dispersion of a given distribution. It ranges between 0.0 and 1.0, 0.0, representing a situation of perfect equality while it is equal to 1.0 in a case of perfect inequality. The Gini is twice the surface between the Lorenz curve and the first diagonal. Gastwirth (1972) has provided a nonparametric method to estimate the Lorenz curve together with a lower bound \( G_L \) and an upper bound \( G_U \) for the Gini coefficient when the data are available as income classes. Schader and Schmid (1994) recommends using \( G_L/3 + 2G_U/3 \) as a point estimate for the Gini coefficient. The method works as follows. Let us suppose that we have \( k \) income classes, the lower bound being \( x_0 = 0 \) and the upper bound \( x_k = +\infty \). The class frequencies are \( n c_i \) with \( n = \sum_{i=1}^{k} n c_i \). The class means are \( \mu_i \) while the overall mean is \( \mu = \sum_{i=1}^{k} \mu_i n c_i / n \). The \( \mu_i \) are estimated using the class boundaries except for the last one which comes from the Pareto assumption \( (\mu_p = \frac{ax_p}{\alpha - 1}) \). Let \((p_i, y_i)\) be the cumulative population share and income share with starting point \((p_0, y_0) = (0, 0)\) and terminal point \((p_k, y_k) = (1, 1)\).

The intermediate points of the Lorenz curve are defined as \( p_i = \sum_{j=1}^{i} n c_j / n, \ y_i = L(p_i) = \sum_{j=1}^{i} \mu_j n c_j / n \) (which are just the generalisation of the natural estimator of the Lorenz curve when raw data are available). The lower bound of the Gini is estimated as:

\[ G_L = 1 - \sum_{i=1}^{k} (y_i + y_{i-1})(p_i - p_{i-1}). \]
This is a lower bound because the underlying Lorenz curve corresponds to a series of linear interpolation segments. The upper bound is obtained as the sum of the lower bound and a factor \( \Delta \) defined as:

\[
\Delta = \frac{1}{\mu} \sum_{i=1}^{k-1} \left( \frac{nc_i}{n} \right)^2 \frac{(\mu_i - x_{i-1})(x_i - \mu_i)}{x_i - x_{i-1}} + \left( \frac{nc_k}{n} \right)^2 (\mu_k - x_{k-1})
\]

Following the Pareto assumption, the mean of the last class is estimated as

\[
\mu_k = x_k \frac{\alpha}{\alpha - 1},
\]

provided that \( \alpha > 1 \). For the other classes, we simply use the mean of a uniform distribution:

\[
\mu_i = \frac{x_{i-1} + x_i}{2}.
\]

Applying this procedure to our 724 school districts, the median value of the Gini coefficient is 0.40, which corresponds to a rather high overall income inequality for 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 an even larger dispersion as it ranges from 0.00 to 0.70. We have shown in Figure 5 that the relation between income inequality and income polarisation is rather weak for the 724 school districts.

5 Testing our Theoretical Model with the Data

Each of the 724 school districts is now equipped with an estimated income distribution with its parameters \( \beta, g \) and eventually \( \alpha \) when there is a Pareto member. Each school district is now characterised by its estimated distribution, a polarisation index and a Gini index. We have all the necessary ingredients to confront our theoretical model to the data. Essentially, the econometric model we use is a two regime switching model. There are strong arguments for adopting a Bayesian approach to make inference in this model. Those arguments are for instance detailed in Xun and Lubrano (2016). Bayesian inference for the two regime switching regression model is explained in Bauwens et al. (1999, Chap. 8). The prior information to be given on the threshold determining the switching is crucial for inference as it conditions the identification of the model.

5.1 Determining the Opting-Out Threshold

We first have to determine the income threshold after which there is a majority of private school enrolment (the opting-out threshold). Does private
school concern only the rich households or is a 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 must be aware that the data represent school districts and not households. So the data concern average figures over a district.

![Figure 8: Rate of public school enrolment](image)

We propose a two regime regression model explaining the rate of public enrolment $\Psi$ where the threshold between the two regressions is the level of mean income. We have chosen the range $[60,150]$ as a prior information when the threshold variable is the mean income divided by 1000 and divided also by the district average number of children per household. Our prior range cover both middle class and the richer class. The posterior results we get are as follows:\(^{15}\)

\[
\begin{align*}
\Psi &= 0.861 - 0.055 g + 0.118 \beta \quad x < \tilde{x} \\
\Psi &= 0.899 - 0.182 g - 0.007 \beta \quad x \geq \tilde{x}
\end{align*}
\]

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

\(^{15}\)Student ratios are given between parenthesis. Average characteristics of the two sub-samples were computed as a byproduct of integration. We provide stars in order to guide the classical reader even if this practice is certainly not Bayesian. The usual significance codes are adopted: *** $P \leq 0.001$, ** $P \leq 0.01$, * $P \leq 0.05$. 

28
observations in the first regime with an average polarisation index of 0.311, a mean Gini of 0.394 and an average public 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 enrolment of 0.77, a mean household income of $356,201.

The opting-out threshold is relatively high as it concerns 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 in the first regime, while inequality is higher in the richer group. Figure 9, which displays the posterior density of the opting-out threshold, shows that the latter is rather well determined.

5.2 Public School Quality

Our Proposition 1 says that, depending on the level of a given $\beta$ (the relative proportion of poor), school public spending depends on polarisation, proxied by the disappearance of the middle class. Differently, de la Croix and Doepke (2009) explain that the level of school public spending is globally 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 $\beta$. For the dependent variable, we have the choice between the Instructional Expenditures per Pupil and the Total Expenditure per Pupil which covers a
broader spectrum of expenditures. These expenditures are financed by tax revenues. We have essentially three sources. The main source should be the local taxes, which are property taxes. The State of California provides a tax revenue states taxes. Finally, the Federal State provides redistribution. The sum of these three tax revenues should cover our Total Expenditure per Pupil. We shall focus on the effect of local revenues and federal revenues.

Following Proposition 1, we expect that, for low values of $\beta$, say $\beta < b_1$, polarisation should have a positive effect on public school spending while for high values of $\beta$, say $\beta > b_2$, polarisation should have a negative sign. For in-between 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 school 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 mean income. However, if we manage to introduce tax revenue, we hope to reduce the uncertainty contained in the theoretical model for $b_1 < \beta < b_2$. So a two regime model should be enough. Inference results are provided in Table 3, using a uniform prior information over $[0.1,0.4]$ for the threshold and a diffuse prior on all other parameters.

The first important point is that whatever the endogenous variable (total versus instructional expenditure), the results are quite similar. The change point between the two regimes corresponds always to a relative proportion of poor households equal to 0.190 which 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 coherent with the opting-out threshold we found in the previous section.

This two regime model says that there are clearly two mechanisms at work explaining school expenditures. First and quite naturally, Federal tax

\[16\] The web site of the National Center for Education Statistics provides the necessary definitions. Instructional Expenditures per Pupil covers mainly wages and activities related to the interaction between teachers and students. Total Expenditure per Pupil includes the previous spending 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 next year to obtain $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 the total expenditure, the gross range is between $4,500 and $18,000 with a median value of $8,000.

\[17\] 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 $\alpha$. As a matter of fact, we use the mean income calculated from the estimated values of $g$, $\beta$ and $\alpha$. 

---

30
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 the average household size) is significant and the same in the two regimes (at least when we consider total expenditures). We predicted a negative impact of polarisation when the proportion of poor households is important. We do find this feature in 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 we defined it, is clearly at least not well represented in this regime. The surprising fact is that the mean Public expenditures per Pupil is not so much 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 we have here the compensating effect of Federal tax revenues.

![Figure 11: Mean income distribution in the two regimes](image)

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 thus corresponds to the centre of the middle class. Inequality is less important in this regime (Gini is 0.407) and plays no significant role for explaining public spending. On the contrary, polarisation is more important (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. Polarisation and inequality provide clearly complementary information.

In Figure 11, we provide the graph of the Mean income distribution in the two regimes. When $\beta$ is greater than the threshold (green curve), 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. When $\beta$ is lower than the threshold, the income distribution is shifted to the right to cover the upper middle class and the rich (red curve). The right tail decays slowly and regularly, contrary to the right tail of the other regime, which explains the difference in polarisation. The crucial difference between the two regimes relies in the importance of the segment around $172,000$, which coincides exactly with our estimated opting-out threshold.

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 pioneeristic model developed by de la Croix and Doepke (2009) 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 concerned an innovative mechanism able to capture the consequences of the complex shape of household income distribution on public policies and school quality.

The first theoretical result is that communities composed by a large presence of low-income compared to high-income families negatively impacts the quality of public school, proxied in the paper by public spending per-pupil. This finding is in line with the empirical evidence for which low quality public schools are mainly concentrated in poor areas. The main contribution is related to the analysis of the effect of income polarisation, measured as the decline of the middle class, on voted public policies at community level. We showed that an increase in income polarisation leads to lower (higher) public school 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 shed light on the relevance of the importance of the income polarisation, other than income inequality, in the analysis of education politics.

We tested our theoretical conclusions using micro-data from California 724 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 school quality. We basically use in this paper a two regime switching regression model which proved to be empirically relevant for implementing two features of our theoretical model. A Bayesian approach helped to overcome the difficulties of the classical approach for this particular model where the usual
asymptotic theory does not apply. We first managed to estimate the opting-out threshold, threshold of household income after which a household prefer to register at a private school. This threshold is fairly high, outside 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. This results solves one of our theoretical interrogation concerning the position of the opting-out threshold. Our second result concerns explaining the level of public school spending, depending on the relative proportion of poor households in a district. For richer districts, we recover the result of de la Croix and Doepke (2009) that inequality has a negative impact on public spending. But for poorer districts, polarisation matters in the same way as inequality did for richer districts. These two mechanism were precisely identified and are completed by the impact of Federal tax revenues as a compensation for poorer districts. Therefore depending on the proportion of poor households relatively to rich households, the key variable is either inequality or polarisation and both have a negative impact. Polarisation and inequality are key complementary phenomenon in order to explain public school quality.

We validated our theoretical model on California data. This could be seen as a limitation. California is a large and populated state. However, it is difficult to speak about the disappearing middle class in California. At most could we speak about the disappearing upper middle class. A more representative state of this phenomenon could be Florida, which is still a large and populated state, but with a lower mean household income and a much lower cost of living.
References


