On the political economy of compulsory education

Alessandro Balestrino^{1,2} · Lisa Grazzini³ · Annalisa Luporini^{2,3}

Received: 13 January 2020 / Accepted: 19 February 2021 / Published online: 24 March 2021
© The Author(s) 2021

Abstract

We consider an economy with two categories of agents: entrepreneurs and workers. In *laissez-faire*, the former gain from having their children educated, while the latter, although they may profit from their own education, have no interest in sending their children to school. We first characterise the preferred education policy-cum-redistributive taxation for the two groups, and find that entrepreneurs favour a compulsory education policy while workers prefer a purely redistributive taxation. Each group would like the policy to be entirely financed by the other group. Then, we introduce a political process with probabilistic voting and verify that an equilibrium with both a compulsory education policy and some redistribution may exist in which the workers are constrained but the entrepreneurs, who benefit from hiring educated workers, are not. The redistribution compensates the workers for being constrained by the education policy.

Keywords Education policy · Redistributive taxation · Probabilistic voting

JEL Classification H42 · H52

We are very grateful to two anonymous referees and to the editor for helpful comments. We also wish to thank the participants to the CESifo Area Conference on Employment and Social Protection 2018, PET 2018, IIPF 2018 and EEA 2018 for fruitful suggestions.

✉ Alessandro Balestrino
alessandro.balestrino@unipi.it

¹ University of Pisa, Pisa, Italy

² CESifo, Munich, Germany

³ University of Florence, Florence, Italy

1 Introduction

It is an historical fact that education policy was conceived in terms of free and mandatory public schooling (financed by public funds) when it was introduced in the West (Germany, France and later UK and US); and free and mandatory schooling is still at the basis of the Western educational systems today. Several motives have been identified for the introduction of compulsory education (Fyfe 2005). In Prussia, where such a system was first introduced in 1763, the protestant religious motive seem to have prevailed (on this see also Becker and Woessmann 2010). In France and Italy compulsory education laws, dating back to 1881 and 1861 respectively, are mainly seen as a part of the construction of a national state (see also Cipolla 1969). In Japan, it was the desire for modernisation that drove the introduction of mandatory schooling after the opening to the West in 1886. Also the UK and the US, by far the most industrialized countries at the time, passed compulsory education laws at the end of the XIX century (1880 in the UK, from 1885 to 1918, depending on the States, in the US); this slight delay might come as a surprise, but a possible reason for it has been identified in the need for cheap child labour—for example, Galor (2006) suggests that education was made compulsory only when a literate workforce was needed because of technological progress. In that case, parents who may profit from their children's labour or contribution to home production (Balestrino et al. 2017; Cigno 2013) may have to be forced to send their kids to school. As far as the US are concerned, Bandiera et al. (2018) also stress a nation-building motive aimed at instilling civic values to migrants during the “Age of Mass Migration” going from 1850 to 1914—an attempt at building social capital, one might say.¹

Initially, the length of mandatory schooling and the enforcement of the attendance prescriptions were relatively limited, especially in South European countries, and generally in the countryside where children were seasonally employed in agriculture. After World War II, there has been a steady increase in the length of compulsory education in Western countries. Murtin and Viarengo (2011) show that there has been a strong convergence in the length of mandatory schooling in fifteen western European countries during the period from 1950 to 2000. At the end of the 1930s, the years of compulsory education ranged from three in Portugal to nine in the UK. After the reforms that occurred in the second half of the XX century, the range was reduced to nine-twelve years. According to Murtin and Viarengo (2011), this convergence is to be traced to the decreasing returns to educational investments, and to the related fact that all countries had reached approximately the same level of profitability. Nowadays, this convergence is further reinforced by globalisation. Higher competitiveness in the global markets can only be faced with a more educated workforce.

In developing countries, however, public and free education is not always guaranteed and thus compulsory education is still an issue today, despite its being

¹ That there is a correlation between education and social and political participation is a well established result. In particular Dee (2004) and Milligan et al. (2004) point out a causal effect from education to participation in voting and civic engagement for the US.

one of the main prerequisites not only for economic development but also for democratization and human rights (as an object in itself and as a primary tool in the fight against child labour). Elementary education should be made compulsory according to art. 26 of Universal Declaration of Human Rights (1948) and such principle has been reaffirmed in a number of conventions and treaties up to Goal 4 of the UNDP Sustainable Development Goals, which calls for achieving inclusive and quality education for all, and more specifically “ensures that all girls and boys complete free primary and secondary schooling by 2030”.

The two motives behind the establishment of compulsory education that we mentioned before, industrialization and generally productivity needs on the one side and nation-building on the other side, are of course not necessarily conflicting with each other. However, when we consider the expansion of compulsory education that took place in Europe after the Second World War, we pointed out above that the productivity motive may have been the main pushing factor.²

For these reasons, we focus on the productivity motive, and therefore build a model in which center stage is taken by the actors of the industrial world, entrepreneurs and workers.³ Also, we adopt a positive, rather than a normative, viewpoint. It is of course always possible to argue in favour of compulsory education in normative terms, e.g. because of horizontal equity requirements (Balestrino et al. 2017). However, normative desirability is not enough to explain why compulsory schooling has become an indispensable part of the modern educational policy package. If we take the political economy view that policies are designed according to voters’ preferences by office- or policy-motivated politicians, then it follows that someone’s interests must be furthered by the presence of a mandatory education period. What we require, therefore, is an argument showing that education policy is likely to be part of a winning policy scheme in a political context.⁴

Specifically, we investigate the question whether there might be a social group who is interested in introducing compulsory schooling as part of the equilibrium

² According to Galor and Moav (2006), the productivity motive was also crucial to favour education reforms in the Western world during the process of the Industrial Revolution when a shift towards more skill-intensive production processes occurred. In particular, they provide evidence suggesting that, during the first phase of the Industrial Revolution, physical capital was the fundamental resource for economic growth while, during the second phase of the Industrial Revolution, the increasing importance of human capital in production induced the capitalist elite to support the provision of public education. Further support to such a view has been recently provided by Squicciarini (2020) for France during the Second Industrial Revolution. In particular, she shows how the type of primary education, and specifically, the introduction of a more technical curriculum was key for promoting a more skilled labour force, and thus for economic development.

³ Main actors may change depending on the historical period. For example, by focussing on the Industrial Revolution period, Galor and Moav (2006) and Galor et al. (2009) point out a different conflict of interests between two types of economic elite: the landowners who were interested in decreasing the mobility of the rural workers and thus were not in favour of their education, and the industrialists who instead needed educated workers and supported educational policies for the masses.

⁴ Our focus in the present paper will be on the compulsory nature of the system. As for its subsidisation, the literature on public finance points out the role that subsidies may play in remedying the distortion in individual educational decisions due to the introduction of income taxation (see e.g. Bovenberg and Jacobs 2005).

policy and is endowed with sufficient political power to actually do so.⁵ In our model, as we said, agents are classified into two occupational groups, entrepreneurs and workers. One of the implications of such a division of society is that entrepreneurs have a stronger interest in education policy than workers. The rationale for this is not that entrepreneurs want their children to be well-educated, because they will tend to provide the required education anyway; the point is that they want *the children of their workers to be educated*, in order to enjoy a better work-force one generation ahead. For this reason, entrepreneurs favour compulsory schooling, financed by the tax system; such a scheme should then prevail at the political equilibrium if the entrepreneurs are able to impose their preferred policy.

Notice also that the fact that in our model both entrepreneurs and workers have a say on education policy through their voting behaviour is consistent with our focus on the productivity factor, which we saw is presumably the main one behind the more recent introductions or expansions of the compulsory education system. Indeed, universal suffrage was not present in the countries where compulsory education was first introduced (see above): due to their low education and income levels, at the time workers did not have the right to vote.⁶

Additionally, we may remark that the phenomenon of “industrial paternalism” provides some indirect evidence of the fact that entrepreneurs have in the course of history cared for their workers’ education. Industrial paternalism entailed, by and large, the provision on the entrepreneurs’ part of basic health and education services to their workers. This covers a period ranging roughly from 1860 to 1950 and is quite common across different places and cultures: we have examples in Europe (France: Reid 1985; Italy: Ciuffetti 2004; Finland: Fellman 2019; UK: Dellheim 1987), in Asia (Japan: Tsutsui 1997), in Africa (Belgian Congo, now Democratic Republic of the Congo: Juif 2019) as well as in the US (Tone 1997). It may be difficult to say what the main rationale for this might have been, possibly a host of different motives (many of which are studied in the references above), but it is undeniable that having a minimally healthy and educated workforce tends to increase productivity – see also fn. 2 and fn. 3. Indeed, it is an established stylised fact that the children of educated parents are more likely to acquire an education (see e.g. Checchi 2005 and the references therein): then, it is plausible to conceive of the entrepreneurs as aware of the potential impact of education on productivity because they see the effects of education on themselves and their children. Instead, workers, not experiencing education first-hand, are less likely to know it as an

⁵ See Bourguignon and Verdier (2000) for a paper showing under which conditions a minority oligarchic group may have an incentive to subsidise the education of the poors who are the majority in the society, and thus to favour the emergence of a middle class as well as a democratic transition. On the other hand, Naito and Nishida (2017) distinguish between primary and higher education and show that different patterns may emerge, depending on the amount of human capital of the median voter. Higher education is supported only if the majority of individuals accumulate sufficient human capital. If this is the case, however, a relatively high amount of public resources may be directed to higher education resulting in persistent high income inequality because the poor do not enrol in university.

⁶ However, as we shall see, we assume that the probability of participation is lower for workers; therefore the historical case could be seen as a limiting one theoretically. Also, it is interesting to note that in several countries mandatory schooling and universal suffrage were simultaneously introduced (Brazil: 1988; South Korea: 1948; India: 1950).

investment for their children and may easily end up in a vicious circle whereby, generation after generation, they prevent their own offspring from reaping the benefits of having an education.⁷ These are the main reasons why we regard the interpretation of the emergence of compulsory education as a response to productivity requirements as more convincing than alternative explanations such as those based on the presence of a majority of poor workers caring for their children's welfare and voting for an education system financed by a tax on rich entrepreneurs. Note, however, that the two explanations are not contradictory and that some altruism on the part of the workers may reinforce the effects of the productivity argument.

The paper is structured as follows. Section 2 presents the model and illustrates the nature of the free-market equilibrium, while Sect. 3 introduces the policy instruments and discusses the policy preferences of the agents. Section 4 expounds the political equilibrium achieved *via* a probabilistic voting process. Finally, Sect. 5 concludes.

2 The model

We consider an overlapping generations economy in which agents live three periods, $i = 0, 1, 2$. In period 0, however, an agent has only a passive role: she receives an education and supplies the time not absorbed by the educational process for the production of a domestically produced service. We refer to agents in period 0 as “children”, in period 1 as “young adults” and in period 2 as “mature adults”. The latter two are the periods where economically relevant decisions are taken and carried out. Agents cease to exist at the end of period 2. For our purposes, then, there are two economically active generations that overlap at each time of the economy, young adults, y , and mature adults, m .

Agents live in households, each made of one parent and one child; in turn, this child will grow up to become a parent; and so on and so forth. There are two social groups, entrepreneurs and workers, who perpetuate themselves generation after generation (no interclass mobility).⁸ Kids are born in period 1, when parents are young adults; in the same period, each parent decides how much education her child should receive. Education requires a money input (out-of-pocket expenditure) as well as a time input (opportunity cost); the time that the kid does not spend in education is combined with the parent's time and employed to provide a household public good. Notice that it is important to characterize the educational process in such a way that the kid's time allocation is explicitly accounted for: indeed, it is exactly because parents may wish to rely on their children's time for the provision

⁷ This may be one of the reasons why the length of compulsory education is found to be a predictor of actual education some 40 years later when the choices of educated parents display their effect on their children's education (Murtin and Viarengo 2011).

⁸ This assumption simplifies our analysis and allows us to focus on the productivity reason behind the political choice on a compulsory educational policy. Should the model be developed in a dynamic set-up, interclass mobility would be crucial to capture the emergence of a more complex social framework (e.g. growth of a middle class, shares in ownership of means of production by workers, etc.).

of the household public good that they may also wish to reduce or ban altogether school's attendance.⁹ This is why we model monetary expenditures and time employment as separate inputs in the children's education.

2.1 Incomes and preferences

There are n entrepreneurs ($n/2$ young and $n/2$ mature adults) resulting in $n/2$ firms. Entrepreneurs' incomes are given by the profits generated by the firms they own. The ownership structure is thus specified: each young adult entrepreneur co-owns the firm with her parent, and they share the profits; one period ahead, the same agent, now a mature adult, will share ownership and earnings with her own child (again, this is just for simplicity, and without loss of generality). Monetary earnings are not the only objective of an entrepreneur who also cares about his reputation as a successful manager of the firm. Since the actions of the entrepreneur display part of their effect after the latter's death, we assume that the entrepreneur will take it into account when making her decisions.

Each firm produces a share of the only good that exists in the economy, whose price is unity. Labour is the only (variable) input and there are constant returns to scale. Each worker supplies a fixed amount of labour, the same for all, and produces $y^i = \underline{y} + y(e_\omega^t, d_\omega^t)$ units of the good in each productive period ($i = 1, 2$), where $\underline{y}$ is the minimum level of production by an uneducated worker, e_ω^t represents the amount of educational expenditure bestowed upon, and d_ω^t denotes the time spent in education by a worker of generation t in period 0 (the total time available is normalized to 1, so that $1 - d_\omega^t$ is the time devoted to the production of the household public good). $y(\cdot)$ is an increasing and strictly concave function satisfying

$$y(0, d_\omega^t) = y(e_\omega^t, 0) = 0; \quad \frac{\partial^2 y}{\partial e_\omega^t \partial d_\omega^t} = \frac{\partial^2 y}{\partial d_\omega^t \partial e_\omega^t} > 0; \quad (1)$$

that is, both inputs into the educational process are essential in order to produce more than the minimum level, $\underline{y}$, and they exhibit technological complementarity: the more time you spend in education, the more effective is the money you spend on it and *viceversa* (for example, if a kid goes to a high-quality school costing more money, this should make the time spent in education more profitable).¹⁰ The agent's non-working time, which is clearly also fixed, is employed in the production of a household public good.

Workers and entrepreneurs bargain over the sharing of output $y^i = \underline{y} + y(e_\omega^t, d_\omega^t)$. Let μ be the index of the power of the entrepreneurs and

⁹ Alternatively, the kids could be employed in a form of market work—the logic of the model would be the same although there would be a further layer of complexity.

¹⁰ The assumption that both time and money are essential to obtain an education outcome is crucial for our analysis. In fact, it would be impossible that one could get educated devoting no time to such activity and using no resources. Complementarity between d and e , on the contrary, is not necessary to obtain our results but, in addition to being a reasonable assumption, it is technically convenient (it simplifies the analysis by ensuring that the agent's problem is well-behaved). Given that both time and money are essential inputs however, we would obtain qualitatively analogous results in case d and e were either independent or substitute in the $y(\cdot)$ function.

$(1 - \mu)$ the index of that of the workers, with $\mu > 0.5$, i.e. entrepreneurs have higher bargaining power than workers. Each firm incurs in bargaining costs that are increasing in the number of workers. We assume that such costs increase in a discontinuous way (for example because at some point a further increase in the number of employees requires an additional person to carry on the bargaining effort) implying that each firm, anticipating correctly its bargaining costs, employs a fixed number of workers that represents the equilibrium profit-maximizing level of employment.¹¹ We denote such level of employment by $2s$, $s \geq 1$, where s is both the number of young adult and that of mature adult workers.¹² Given that there are $n/2$ firms we globally have $S = sn \geq n$ employed workers.

The objective of the entrepreneurs is to maximize their share of per-period profit $\underline{y} + y(e_\omega^t, d_\omega^t) - w_\omega^{it} - C$ where w_ω^{it} represents the per-period wage, and C is the per-worker bargaining cost; while the objective of a worker is to maximize w_ω^{it} . We assume (generalized) Nash bargaining and posit i) that the bargaining costs are sunk, because they are borne independently of the bargaining outcome and ii) that bargaining occurs before production costs are incurred: then, the disagreement point is $(-C, 0)$ ¹³ and the wage level will result from the solution of

$$\max_{w_\omega^{it}} (\underline{y} + y(e_\omega^t, d_\omega^t) - w_\omega^{it})^\mu (w_\omega^{it})^{1-\mu},$$

which yields

$$w_\omega^{it} = (1 - \mu)[\underline{y} + y(e_\omega^t, d_\omega^t)]. \quad (2)$$

Assuming a perfect credit market with zero interest rate, a worker thus earns lifetime income

$$w_\omega^t = 2(1 - \mu)[\underline{y} + y(e_\omega^t, d_\omega^t)]. \quad (3)$$

The per-worker profit in each period will be

$$\pi^{i,t}(e_\omega^t, d_\omega^t) = \underline{y} + y(e_\omega^t, d_\omega^t) - w_\omega^{i,t}(e_\omega^t, d_\omega^t) = \mu[\underline{y} + y(e_\omega^t, d_\omega^t)] \quad i = 1, 2. \quad (4)$$

Each entrepreneur earns lifetime income

$$\begin{aligned} w_\eta^t = & w_\eta^{1,t}(e_\omega^{t-1}, d_\omega^{t-1}, e_\eta^{t-1}, d_\eta^{t-1}, e_\omega^t, d_\omega^t, e_\eta^t, d_\eta^t) \\ & + w_\eta^{2,t}(e_\omega^t, d_\omega^t, e_\eta^t, d_\eta^t, e_\omega^{t+1}, d_\omega^{t+1}, e_\eta^{t+1}, d_\eta^{t+1}), \end{aligned} \quad (5)$$

¹¹ In other words, we assume that the marginal bargaining cost in equilibrium is higher than the profit obtainable by employing an additional worker.

¹² This is of course just an innocuous simplification; the model works with any share of young to mature adults employed in each firm.

¹³ This can be interpreted as meaning that the bargaining costs of the worker are normalised to zero.

where the subscript η denotes a variable pertaining to an entrepreneur. The lifetime income is given by the sum of the entrepreneur incomes in period 1 and 2, $w_\eta^{1,t}$ and $w_\eta^{2,t}$, where

$$w_\eta^{1,t} \left(e_\omega^{t-1}, d_\omega^{t-1}, e_\eta^{t-1}, d_\eta^{t-1}, e_\omega^t, d_\omega^t, e_\eta^t, d_\eta^t \right) \\ = \alpha \left\{ \left[\pi^{1,t} (e_\omega^t, d_\omega^t) + \pi^{1,t-1} (e_\omega^{t-1}, d_\omega^{t-1}) \right] s + g(e_\eta^{t-1}, d_\eta^{t-1}) + g(e_\eta^t, d_\eta^t) \right\}; \quad (6)$$

$$w_\eta^{2,t} \left(e_\omega^t, d_\omega^t, e_\eta^t, d_\eta^t, e_\omega^{t+1}, d_\omega^{t+1}, e_\eta^{t+1}, d_\eta^{t+1} \right) \\ = (1 - \alpha) \left\{ \left[\pi^{2,t+1} (e_\omega^{t+1}, d_\omega^{t+1}) + \pi^{2,t} (e_\omega^t, d_\omega^t) \right] s + g(e_\eta^t, d_\eta^t) + g(e_\eta^{t+1}, d_\eta^{t+1}) \right\}, \quad (7)$$

and where $\alpha(1 - \alpha)$ is the share of earning accruing to a young (mature) adult; $g(\cdot)$ is an increasing and concave function converting, for both co-owners of the firm and in each period, the educational inputs received into income—such a function might therefore represent the returns to entrepreneurial ability as mediated by the investments in human capital. Mirroring the preceding assumptions on $y(\cdot)$, we posit

$$g(0, d_\eta^t) = g(e_\eta^t, 0) = 0; \quad \frac{\partial^2 g}{\partial e_\eta^t \partial d_\eta^t} = \frac{\partial^2 g}{\partial d_\eta^t \partial e_\eta^t} > 0, \quad (8)$$

so that both educational expenditure and time spent in education are essential to develop entrepreneurial ability and the two inputs exhibit technological complementarity.

Also mirroring the assumptions made on the workers' time allocation, we assume that each entrepreneur supplies a fixed amount of time for management, the same for all, and that the remaining fixed leisure time is employed along with the kid's non-educational time to produce a household public good.

As a final remark, we notice that, given $\mu > 0.5$, (3) and (5) imply that the worker's lifetime income is lower than the entrepreneur's lifetime income:

$$w_\omega^t < w_\eta^t, \quad \forall t. \quad (9)$$

Turning now to the agents' preferences, we assume that all agents are selfish.¹⁴ Neither the workers nor the entrepreneurs derive any direct utility from their children's education. However, while workers do not derive any indirect utility either, young entrepreneurs derive an indirect advantage from investing in their own children's education because this positively affects next-period profits—see (7). Moreover, mature entrepreneurs exhibit a concern for their firm's future

¹⁴ This is of course an extreme assumption, and we adopt it mainly because it simplifies the reasoning quite radically. It has to be remembered, however, that while the verdict on whether altruism or egoism prevails within a family is possibly still open, the assumption of egoism seems to be more robust to empirical scrutiny (see e.g. Cigno et al. 1998, 2006). Moreover, our results would carry over to a setting with altruism, as long as the market equilibrium yields a less-than-optimal amount of acquired education, and Balestrino et al. (2017) show that even with full altruism there can be inefficiency in the provision of education due to comparative advantage issues.

profitability, a “reputational effect”. We choose to focus on this specific element due to our interest in the productivity/industrialisation rationale for educational policy: to keep things simple, we ignore other possible variants like assuming that the parents take pride in their offspring’s education or are worried about the effect of the children’s homework. The impact of the reputational effect on the agents’ voting preferences is discussed in Sect. 4, where it is made clear that the introduction of such an effect brings about only qualitative changes in the results and that nothing of substance is modified.

Therefore, the workers’ utility function is

$$U_{\omega} = u(c_{\omega}^{1,t}) + v(c_{\omega}^{2,t}) + f(1 - d_{\omega}^{t+1}), \quad (10)$$

where $f(\cdot)$ represents the utility from the production of the household public good that we mentioned above. Since the parent’s leisure is fixed, however, we write the sub-utility directly as a function of the kid’s domestic time only, with the provision that $f(0) > 0$ —i.e. that only parental time is essential to the production of the household public good.

As to the utility function of the entrepreneurs, it still depends on consumption and on the provision of the household public good. Moreover, we capture the reputational effect by directly introducing a fraction β , $0 < \beta < 1$, of future profits in the utility function:

$$\begin{aligned} U_{\eta} = & u(c_{\eta}^{1,t}) + v(c_{\eta}^{2,t}) + f(1 - d_{\eta}^{t+1}) \\ & + \beta w_{\eta}^{3,t}(e_{\omega}^{t+1}, d_{\omega}^{t+1}, e_{\eta}^{t+1}, d_{\eta}^{t+1}, e_{\omega}^{t+2}, d_{\omega}^{t+2}, e_{\eta}^{t+2}, d_{\eta}^{t+2}) \end{aligned} \quad (11)$$

where

$$\begin{aligned} w_{\eta}^{3,t}(e_{\omega}^{t+1}, d_{\omega}^{t+1}, e_{\eta}^{t+1}, d_{\eta}^{t+1}, e_{\omega}^{t+2}, d_{\omega}^{t+2}, e_{\eta}^{t+2}, d_{\eta}^{t+2}) \\ = [\pi^{3,t+1}(e_{\omega}^{t+1}, d_{\omega}^{t+1}) + \pi^{3,t+2}(e_{\omega}^{t+2}, d_{\omega}^{t+2})]s \\ + g(e_{\eta}^{t+1}, d_{\eta}^{t+1}) + g(e_{\eta}^{t+2}, d_{\eta}^{t+2}), \end{aligned} \quad (12)$$

is the profit generated in the period following the death of the entrepreneur.

We start by describing the *laissez-faire* economy; government interventions will be considered later on.

2.2 Agent optimisation in a free market

Each worker maximises (10) by choosing her consumption basket and the composition of her kid’s educational process subject to her lifetime budget constraint

$$c_{\omega}^{1,t} + c_{\omega}^{2,t} + e_{\omega}^{t+1} = 2(1 - \mu)[\underline{y} + y(e_{\omega}^t, d_{\omega}^t)] \quad (13)$$

and her child time constraint

$$d_{\omega}^{t+1} \leq 1, \quad (14)$$

plus non-negativity constraints for all the choice variables. Since the educational expenditure for the next generation e_{ω}^{t+1} does not appear in the utility function, and the time spent by children in education d_{ω}^{t+1} appears as a bad, it is clear that $e_{\omega}^{t+1} = d_{\omega}^{t+1} = 0$ at the optimum for all workers of all generations. Thus, the problem reduces to

$$\text{Max } u(c_{\omega}^{1,t}) + v(2(1 - \mu)y - c_{\omega}^{1,t}), \quad (15)$$

where the budget constraint (13) has been substituted into the utility function. The FOC w.r.t. $c_{\omega}^{1,t}$ is, quite simply,

$$u' = v'. \quad (16)$$

Workers smooth their consumption over time. Since no worker gains from sending her child to school, however, the workers never get an education.

As for the entrepreneurs, they maximise (11) subject to their lifetime budget constraint

$$\begin{aligned} c_{\eta}^{1,t} + c_{\eta}^{2,t} + e_{\eta}^{t+1} &= w_{\eta}^{1,t} \left(e_{\omega}^{t-1}, d_{\omega}^{t-1}, e_{\eta}^{t-1}, d_{\eta}^{t-1}, e_{\omega}^t, d_{\omega}^t, e_{\eta}^t, d_{\eta}^t \right) \\ &+ w_{\eta}^{2,t} \left(e_{\omega}^t, d_{\omega}^t, e_{\eta}^t, d_{\eta}^t, e_{\omega}^{t+1}, d_{\omega}^{t+1}, e_{\eta}^{t+1}, d_{\eta}^{t+1} \right) \end{aligned} \quad (17)$$

and their time constraint

$$d_{\eta}^{t+1} \leq 1. \quad (18)$$

Letting λ denote the Lagrange multiplier for the budget constraint, the FOCs w.r.t. $c_{\eta}^{1,t}$, $c_{\eta}^{2,t}$, e_{η}^{t+1} , and d_{η}^{t+1} are

$$u' = \lambda; v' = \lambda; \quad \lambda \left[(1 - \alpha) \frac{\partial g}{\partial e_{\eta}^{t+1}} - 1 \right] + \beta \frac{\partial g}{\partial e_{\eta}^{t+1}} = 0; \quad [\lambda(1 - \alpha) + \beta] \frac{\partial g}{\partial d_{\eta}^{t+1}} \geq f', \quad (19)$$

respectively, so that the budget allocation is ruled by

$$u' = v' = [\lambda(1 - \alpha) + \beta] \frac{\partial g}{\partial e_{\eta}^{t+1}}. \quad (20)$$

Again, consumption will be smoothed over the two periods; however, as far as the entrepreneurs are concerned, each of them gains from having her kid educated, because in the next two periods that kid will own part of the firm, and will contribute her managerial skills to the production process and thus first to the earnings and then to the reputation of the entrepreneur. Therefore, children belonging to this class are educated, and might indeed go to school full-time ($d_{\eta}^{t+1} = 1$). Notice that the

reputational effect ($\beta > 0$) raises the levels of e_{η}^{t+1} and d_{η}^{t+1} but is not necessary for the entrepreneurs to educate their children.

2.3 Characteristics of the free market equilibrium

In the *laissez-faire* equilibrium, some agents (the entrepreneurs) educate their children while others (the workers) don't. Notice that the reason why workers are not educated is that educational expenses must be paid by the parent, but the latter does not obtain any return from her child's education. *Not only, but the time devoted to education is subtracted from the production of the household public good.* The entrepreneurs, on the contrary, in addition to the gain they get from educating their children, may also take advantage from having an educated work force. This may open the way for policies that oblige parents to send their kids to school.

3 Agent optimisation and policy preferences

In order to investigate whether a compulsory public education policy could gain the support of the majority of voters, we must first assess whether such a measure can actually improve the welfare either of the entrepreneurs, or of the workers or of both categories. As far as the policy tools are concerned, we consider a compulsory education package and a linear income tax/subsidy to be employed both for financing such education measures and for redistributive purposes. We let τ_{ω} and τ_{η} denote the group-specific marginal income tax rates for workers and entrepreneurs, respectively (possibly negative),¹⁵ while $\bar{e}$ represents the minimum expenditure on a child's education that is imposed upon households and $\bar{d}$ the minimum amount of time that a child must spend in school. Consequently e and d will now represent the amounts of money and time that are freely allocated to education by households on top of the prescribed levels. Notice that, since the time allocation for the parent is fixed, τ_{ω} and τ_{η} are not distortionary, and basically equivalent to lump-sum transfers.

3.1 Agent optimisation in the presence of an active policy

Let's take the workers. Taking into account (2) and the education policy described above, a worker per-period after tax income obtains as

$$(1 - \tau_{\omega})w_{\omega}^{i,t} \equiv (1 - \tau_{\omega})(1 - \mu)\left(\underline{y} + y(\bar{e} + e_{\omega}^t, \bar{d} + d_{\omega}^t)\right). \quad (21)$$

¹⁵ The model cannot be applied to the case of a single income tax rate, because, as we shall see, the key mechanism driving the agents' policy preferences is the desire of each group to shift the tax burden on the other group while retaining the benefit of the public expenditure. However, the case we consider here is more relevant, we believe, because, as the gross incomes of the two groups are different (see (9)), it becomes a rough representation of a two-bracket income tax system - therefore, it approximates more closely real-world tax systems.

Further, the worker budget constraint (13) becomes

$$c_{\omega}^{1,t} + c_{\omega}^{2,t} + \bar{e} + e_{\omega}^{t+1} = 2(1 - \tau_{\omega}) \left(\underline{y} + y(\bar{e} + e_{\omega}^t, \bar{d} + d_{\omega}^t) \right) + \bar{e}, \quad (22)$$

where $\bar{e}$ appears on both sides of the constraint because the policy is publicly financed (either each household receives a subsidy or monetary educational expenses are paid by the government). The time constraint of worker's child (14) becomes

$$d_{\omega}^{t+1} \leq 1 - \bar{d}. \quad (23)$$

Just as in the free-market equilibrium, the additional education expenditure for children e_{ω}^{t+1} does not appear in the utility function, and the additional time spent by children in education d_{ω}^{t+1} appears as a bad, therefore $e_{\omega}^{t+1} = d_{\omega}^{t+1} = 0$ at the optimum for all workers of all generations. Thus, a worker's maximization problem reduces to

$$\text{Max } u(c_{\omega}^{1,t}) + v \left(2(1 - \tau_{\omega})(1 - \mu) \left(\underline{y} + y(\bar{e}, \bar{d}) \right) - c_{\omega}^{1,t} \right) + f(1 - \bar{d}). \quad (24)$$

The FOC w.r.t. $c_{\omega}^{1,t}$ obtains as

$$u' = v', \quad (25)$$

i.e. it has the same form as the FOC (16) obtained in a free market, leading again to consumption smoothing. But now the worker is obliged to have the kid spend $\bar{d}$ as study time. The worker will also spend $\bar{e}$ on her child's education but this would be financed by the tax system.

Let us now consider the entrepreneurs: the entrepreneur's budget constraint (17) becomes

$$c_{\eta}^{1,t} + c_{\eta}^{2,t} + \bar{e} + e_{\eta}^{t+1} = (1 - \tau_{\eta}) \left(w_{\eta}^{1,t} + w_{\eta}^{2,t} \right) + \bar{e}. \quad (26)$$

Since $e_{\omega}^t = d_{\omega}^t = 0$ for the reasons given above, we have from (6) and (7) that the entrepreneur's income in period 1, 2, and 3 obtain as follows

$$\begin{aligned} & w_{\eta}^{1,t} \left(\bar{e}, \bar{d}, e_{\eta}^{t-1}, d_{\eta}^{t-1}, e_{\eta}^t, d_{\eta}^t \right) \\ &= \alpha \left\{ \left[\pi^{1,t}(\bar{e}, \bar{d}) + \pi^{1,t-1}(\bar{e}, \bar{d}) \right] s + g \left(\bar{e} + e_{\eta}^{t-1}, \bar{d} + d_{\eta}^{t-1} \right) \right. \\ & \quad \left. + g \left(\bar{e} + e_{\eta}^t, \bar{d} + d_{\eta}^t \right) \right\}; \end{aligned} \quad (27)$$

$$\begin{aligned} & w_{\eta}^{2,t} \left(\bar{e}, \bar{d}, e_{\eta}^t, d_{\eta}^t, e_{\eta}^{t+1}, d_{\eta}^{t+1} \right) \\ &= (1 - \alpha) \left\{ \left[\pi^{2,t+1}(\bar{e}, \bar{d}) + \pi^{2,t}(\bar{e}, \bar{d}) \right] s + g \left(\bar{e} + e_{\eta}^t, \bar{d} + d_{\eta}^t \right) \right. \\ & \quad \left. + g \left(\bar{e} + e_{\eta}^{t+1}, \bar{d} + d_{\eta}^{t+1} \right) \right\}; \end{aligned} \quad (28)$$

$$\begin{aligned}
 w_{\eta}^{3,t}(\bar{e}, \bar{d}, e_{\eta}^{t+1}, d_{\eta}^{t+1}, e_{\eta}^{t+2}, d_{\eta}^{t+2}) \\
 = [\pi^{3,t+1}(\bar{e}, \bar{d}) + \pi^{3,t+2}(\bar{e}, \bar{d})]s + \left[g(e_{\eta}^{t+1}, d_{\eta}^{t+1}) + g(e_{\eta}^{t+2}, d_{\eta}^{t+2}) \right].
 \end{aligned} \quad (29)$$

Entrepreneurs maximise their utility function

$$U_{\eta} = u(c_{\eta}^{1,t}) + v(c_{\eta}^{2,t}) + f(1 - \bar{d} - d_{\eta}^{t+1}) + \beta w_{\eta}^{3,t} \quad (30)$$

by choice of $c_{\eta}^{1,t}$, $c_{\eta}^{2,t}$, e_{η}^{t+1} and d_{η}^{t+1} subject to the budget constraint (26) and the additional time constraint

$$d_{\eta}^{t+1} \leq 1 - \bar{d}. \quad (31)$$

Since it will become clear in the next subsection that there cannot exist a political equilibrium where both entrepreneurs and workers are constrained, we only consider interior solutions for e_{η}^{t+1} and d_{η}^{t+1} . The FOCs then are

$$\begin{aligned}
 u' = \lambda; \quad v' = \lambda; \quad \lambda \left[(1 - \tau_{\eta})(1 - \alpha) \frac{\partial g}{\partial e_{\eta}^{t+1}} - 1 \right] + \beta \frac{\partial g}{\partial e_{\eta}^{t+1}} = 0; \\
 [\lambda(1 - \tau_{\eta})(1 - \alpha) + \beta] \frac{\partial g}{\partial d_{\eta}^{t+1}} = f'.
 \end{aligned} \quad (32)$$

3.2 Policy preferences

We now have to check which of the possible constellations of policy tools is preferred by the agents. Let us begin by writing the government revenue constraint under the assumption that the educational expenditure ration $\bar{e}$ is paid for by the government:

$$\tau_{\omega}(1 - \mu) \left(\underline{y} + y^{t-1} + \underline{y} + y^t \right) \frac{S}{2} + \tau_{\eta} \left(w_{\eta}^{1,t} + w_{\eta}^{2,t-1} \right) \frac{n}{2} = \frac{(n + S)}{2} \bar{e}, \quad (33)$$

where we dropped the arguments in y^t , y^{t-1} , $w_{\eta}^{1,t}$ and $w_{\eta}^{2,t-1}$ to avoid clutter. For future use, we write the public budget in *per-capita* terms and we express it in terms of $\tau_{\omega}(\tau_{\eta}, \bar{e}, \bar{d})$:

$$\tau_{\omega}(\tau_{\eta}, \bar{e}, \bar{d}) = \frac{\bar{e} - \tau_{\eta} \left(w_{\eta}^{1,t} + w_{\eta}^{2,t-1} \right) (1 - \sigma)}{\Psi \sigma}, \quad (34)$$

where $\sigma = S/(n + S)$ and $\Psi = (1 - \mu) \left(2\underline{y} + y^t + y^{t-1} \right)$. Next, by deriving $\tau_{\omega}(\cdot)$ with respect to τ_{η} , $\bar{e}$, and $\bar{d}$ we obtain

$$\frac{\partial \tau_{\omega}}{\partial \tau_{\eta}} = - \frac{1 - \sigma}{\Psi \sigma} < 0, \quad (35)$$

$$\frac{\partial \tau_{\omega}}{\partial \bar{e}} = \frac{1}{\Psi \sigma} > 0, \quad (36)$$

$$\frac{\partial \tau_{\omega}}{\partial \bar{d}} = \frac{-\tau_{\eta} \left(\frac{\partial w_{\eta}^{1,t}}{\partial \bar{d}} + \frac{\partial w_{\eta}^{2,t-1}}{\partial \bar{d}} \right) (1 - \sigma) \Psi \sigma - 2(1 - \mu) \frac{\partial y}{\partial \bar{d}} \sigma (\bar{e} - \tau_{\eta} (w_{\eta}^{1,t} + w_{\eta}^{2,t-1}) (1 - \sigma))}{(\Psi \sigma)^2}. \quad (37)$$

Notice that here the production function is represented as affected by the education level of the *parents*, y^t , and the *grandparents*, y^{t-1} , while the possible increase in education prescribed by the policy would affect the earnings of the *children*. Similarly the current revenue of the entrepreneurs, $w_{\eta}^{1,t}$ and $w_{\eta}^{2,t-1}$, is not affected by a change in $\bar{d}$. This implies the following

$$\frac{\partial y^{t-1}}{\partial \bar{d}} = \frac{\partial y^t}{\partial \bar{d}} = \frac{\partial w_{\eta}^{1,t}}{\partial \bar{d}} = \frac{\partial w_{\eta}^{2,t-1}}{\partial \bar{d}} = 0. \quad (38)$$

By using (38), (37) can be re-written as

$$\frac{\partial \tau_{\omega}}{\partial \bar{d}} = 0. \quad (39)$$

Let the indirect utility, written as a function of the policy instruments, be denoted by

$$V_i = V_i(\tau_i, \bar{e}, \bar{d}), \quad i = \omega, \eta_y, \eta_m, \quad (40)$$

where η_y denotes a young entrepreneur and η_m denotes a mature entrepreneur. The derivatives of (40) with respect to the policy instruments for the workers are

$$\frac{\partial V_{\omega}}{\partial \tau_{\omega}} = -2(1 - \mu) (\underline{y} + y^t) v' < 0; \quad (41)$$

$$\frac{\partial V_{\omega}}{\partial \bar{e}} = (1 - \tau_{\omega}) 2(1 - \mu) \frac{\partial y^t}{\partial \bar{e}} v'; \quad (42)$$

$$\frac{\partial V_{\omega}}{\partial \bar{d}} = \left((1 - \tau_{\omega}) 2(1 - \mu) \frac{\partial y^t}{\partial \bar{d}} \right) v' - f'; \quad (43)$$

where again,

$$\frac{\partial y^t}{\partial \bar{e}} = \frac{\partial y^t}{\partial \bar{d}} = 0, \quad (44)$$

as far as the parents' and grandparents' income is concerned. Therefore, (42) and (43) can be re-written as

$$\frac{\partial V_{\omega}}{\partial \bar{e}} = 0; \quad \frac{\partial V_{\omega}}{\partial \bar{d}} = -f' < 0. \quad (45)$$

Regarding the entrepreneurs, we must distinguish between the young and the mature ones. For the *young*, the derivative of (40) with respect to the entrepreneur's income tax rate obtains as

$$\frac{\partial V_{\eta_y}}{\partial \tau_{\eta}} = -\left(w_{\eta}^{1,t} + w_{\eta}^{2,t}\right)\lambda < 0. \quad (46)$$

As to the derivatives with respect to the minimum expenditure on a child's education, $\bar{e}$, and the minimum amount of time a child must spend in school, $\bar{d}$, we have

$$\frac{\partial V_{\eta_y}}{\partial \bar{e}} = (1 - \tau_{\eta}) \frac{\partial w_{\eta}^{2,t}}{\partial \bar{e}} \lambda + \beta \frac{\partial w_{\eta}^{3,t}}{\partial \bar{e}} > 0; \quad \frac{\partial V_{\eta_y}}{\partial \bar{d}} = (1 - \tau_{\eta}) \frac{\partial w_{\eta}^{2,t}}{\partial \bar{d}} \lambda + \beta \frac{\partial w_{\eta}^{3,t}}{\partial \bar{d}} > 0, \quad (47)$$

where we have considered that

$$\frac{\partial w_{\eta}^{1,t}}{\partial \bar{e}} = \frac{\partial w_{\eta}^{1,t}}{\partial \bar{d}} = 0, \quad (48)$$

because education affects only next-period profits. Notice that the per period entrepreneurs' income is made of four terms – see (27) and (28). Since the entrepreneurs are not constrained, the compulsory education policy does not induce any change in returns to period two and three entrepreneurial activity $(1 - \alpha)g$, but, given (4), it creates more income *via* increases in per-worker profits $\pi^{2,t}$ and $(1 - \tau_{\eta})\pi^{3,t}$. This means that we can be certain that

$$\frac{\partial w_{\eta}^{2,t}}{\partial \bar{e}} > 0; \frac{\partial w_{\eta}^{2,t}}{\partial \bar{d}} > 0; \frac{\partial w_{\eta}^{3,t}}{\partial \bar{e}} > 0; \frac{\partial w_{\eta}^{3,t}}{\partial \bar{d}} > 0. \quad (49)$$

Consequently the sign of the derivatives of (47) is positive. In fact the policy measure has no impact on the amount of time and money invested in the education of an entrepreneur's child. The increase in the compulsory components of e and d will be counterbalanced by a reduction of the same amount in the time and money used to top up the compulsory amounts. As a consequence the entrepreneurs will benefit from the increase of the education of their work-force without incurring in any distortion of their own educational choices.

The *mature* entrepreneurs will incur in the cost of education without obtaining any monetary return, but obtaining instead a benefit in terms of reputation. The reputational effect, therefore, is key to make them favour an educational policy, in that it makes them care, indirectly, for the workers' children's education one generation ahead. Qualitatively, this works just as if we assumed that the mature entrepreneurs had become altruistic (which, as we mentioned in fn. 14 would still be compatible with our results); the interpretation as an interest in the reputation of the firm is however more in line with our basic assumptions. While helpful, it is not absolutely necessary to obtain our main result (the interest in workers' education of the young entrepreneurs would be enough), and it allows us to show the entrepreneurs as being consistently in support of mandatory education rather than

moving away from that support in old age, which seems to sit comfortably with a view of the entrepreneur as having a long-term interest in the family business.

For them, the derivatives of (40) with respect to τ_η , $\bar{e}$, and $\bar{d}$ obtain as

$$\frac{\partial V_{\eta_m}}{\partial \tau_\eta} = -\left(w_\eta^{2,t-1}\right)\lambda < 0; \quad \frac{\partial V_{\eta_m}}{\partial \bar{e}} = \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{e}}; \quad \frac{\partial V_{\eta_m}}{\partial \bar{d}} = \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{d}}. \quad (50)$$

We are now in a position to calculate the preferred policy by each group. Specifically, the preferred policies can be found by using (34) to replace τ_ω in (40) and then choosing τ_η , $\bar{e}$ and $\bar{d}$ so as to maximise:

$$V_\omega = V_\omega((\tau_\omega(\tau_\eta, \bar{e}, \bar{d}), \bar{e}, \bar{d}), \quad (51)$$

$$V_{\eta_k} = V_{\eta_k}(\tau_\eta, \bar{e}, \bar{d}); \quad k = y, m, \quad (52)$$

for the workers and the entrepreneurs, respectively, under non-negativity constraints for $\bar{e}$ and $\bar{d}$ and the constraints that

$$\tau_l \leq 1, \quad l = \omega, \eta; \quad \bar{d} \leq 1. \quad (53)$$

For the workers, the FOCs are:

$$\frac{dV_\omega}{d\tau_\eta} = \frac{\partial V_\omega}{\partial \tau_\omega} \frac{\partial \tau_\omega}{\partial \tau_\eta} = 2(1-\mu)(\underline{y} + y^t) v' \frac{(1-\sigma)}{2(1-\mu)(\underline{y} + y^t)\sigma} = v' \frac{(1-\sigma)}{\sigma} > 0; \quad (54)$$

$$\frac{dV_\omega}{d\bar{e}} = \frac{\partial V_\omega}{\partial \tau_\omega} \frac{\partial \tau_\omega}{\partial \bar{e}} = -2(1-\mu)(\underline{y} + y^t) v' \frac{1}{2(1-\mu)(\underline{y} + y^t)\sigma} = -\frac{1}{\sigma} v' < 0; \quad (55)$$

$$\frac{dV_\omega}{d\bar{d}} = -f' + \frac{\partial V_\omega}{\partial \tau_\omega} \frac{\partial \tau_\omega}{\partial \bar{d}} = -f' < 0, \quad (56)$$

implying that the optimal tax rate is $\tau_\eta = 1$ while $\bar{e}$ and $\bar{d}$ should be optimally set to zero.

For the *young* entrepreneurs, the FOCs are:

$$\frac{\partial V_{\eta_y}}{\partial \tau_\eta} = -\left(w_\eta^{1,t} + w_\eta^{2,t}\right)\lambda < 0; \quad (57)$$

$$\frac{\partial V_{\eta_y}}{\partial \bar{e}} = (1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{e}} + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{e}} > 0; \quad (58)$$

$$\frac{\partial V_{\eta_y}}{\partial \bar{d}} = (1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{d}} \lambda + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{d}} > 0. \quad (59)$$

The FOCs for the *mature* ones are instead:

$$\frac{\partial V_{\eta_m}}{\partial \tau_\eta} = - \left(w_\eta^{2,t-1} \right) \lambda < 0; \quad (60)$$

$$\frac{\partial V_{\eta_m}}{\partial \bar{e}} = \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{e}} > 0; \quad (61)$$

$$\frac{\partial V_{\eta_m}}{\partial \bar{d}} = \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{d}} > 0. \quad (62)$$

We know from our previous analysis that in this case both $\partial V_{\eta_y} / \partial \bar{e}$ and $\partial V_{\eta_y} / \partial \bar{d}$ are positive because of a positive indirect effect as the compulsory education policy creates more income *via* increases in the after-tax per-worker next-period profits $(1 - \tau_\eta) \pi^{2,t}$ —see (47)—and because of the positive reputational effect. The latter is also present in the case of the mature entrepreneurs.

Therefore, the entrepreneurs would prefer to face a zero marginal tax rate while at the same time having positive values for $\bar{e}$ and $\bar{d}$ (indeed, entrepreneurs would always favour pushing each ration to its upper limit). This implies that the workers should face a positive tax rate in order to finance education expenditure. The upper limit for $\bar{d}$ is clearly unity, while for $\bar{e}$ can be deduced from observing that, given the preferred tax rates, the maximum level of $\bar{e}$ can be achieved when $\tau_\omega(\tau_\eta, \bar{e}, \bar{d}) = 1$, implying $\bar{e} = (1 - \mu) \left(2\underline{y} + y^t + y^{t-1} \right) \sigma$.

While the results are possibly too sharp to be taken literally, their qualitative interpretation is clear: the workers do not perceive any benefit from compulsory education but would favour a redistributive income taxation, whereas entrepreneurs gain from compulsory education but would like to shift the entire cost on the workers.

4 Political equilibrium

Let us now focus on the voting process through which an educational policy package is chosen in the political arena. To perform our analysis, we consider a probabilistic voting model with a two-candidate electoral competition—see e.g. Lindbeck and Weibull (1987). In this setup, candidates are uncertain on whether citizens will participate in voting: they could abstain, maybe because they cannot clearly perceive the distance between the proposed platforms. Consequently, the candidates are uncertain on how citizens will vote for any given political proposal. Following a standard approach, we suppose that the voters' decisions depend on the differences in the expected utilities from the candidates' different platforms, and that the candidates perceive the probability that a voter will participate in voting and support a platform as a function of the distance between her own platform and that proposed by the rival candidate. Politicians are assumed to be opportunistic, i.e. they

are purely office-motivated, and thus aim at maximising their vote share. No credibility issues may arise, because it is also assumed that politicians can make binding commitments to policy platforms proposed during the electoral campaign.¹⁶

To sum up, the sketch of the electoral procedure is thus the following. Two candidates simultaneously propose their policy platforms, that is their educational policy packages plus their redistributive policy platform. Then, citizens vote for their preferred candidate. Finally, the elected candidate implements the policy she promised during the electoral campaign.

Each candidate selects her policy platform in order to maximise her share of total votes, that depends on the probabilities that each voter will vote for her, taking the rival candidate platform as given. Now, let the probability perceived by candidate $j, j = A, B$ that an agent votes for her be $\gamma_i^j, i = \omega, \eta_y, \eta_m$, where we distinguish between young and mature entrepreneurs because they have different policy preferences.¹⁷ The expected vote share of a candidate will then be:

$$p^j = \sigma \gamma_{\omega}^j + \frac{(1 - \sigma)}{2} (\gamma_{\eta_y}^j + \gamma_{\eta_m}^j), \quad j = A, B. \quad (63)$$

As usual, we posit

$$\begin{aligned} \gamma_i^j = & \Gamma_i \left[V_i \left(\tau_{\eta}^j, \tau_{\omega}^j \left(\tau_{\eta}^j, \bar{e}^j, \bar{d}^j \right), \bar{e}^j, \bar{d}^j \right) \right. \\ & \left. - V_i \left(\tau_{\eta}^{-j}, \tau_{\omega}^{-j} \left(\tau_{\eta}^{-j}, \bar{e}^{-j}, \bar{d}^{-j} \right), \bar{e}^{-j}, \bar{d}^{-j} \right) \right], \\ & i = \omega, \eta_y, \eta_m; \quad j = A, B, \end{aligned} \quad (64)$$

where Γ_i is a smooth, continuous and increasing function varying between 0 and 1, and we use the superscript $j, j = A, B$, to denote a policy variable proposed by candidate j .

The assumption that agents will show up at elections with a certain positive probability is of course standard in probabilistic voting models; also standard is it to assume that this probability varies with the agent's type and, more precisely, that each individual's voting behaviour is affected by her own ideological attachment to a party (usually represented by an idiosyncratic taste shock which is a random variable with a density function taken to be symmetric around zero). However, we wish to highlight here a different mechanism, namely the positive relationship between income and voting participation: active participation in public life, including active voting, is indeed usually found to be positively related to income at the individual level and, relatedly, negatively associated with income inequality at

¹⁶ Notice that, since politicians are office- rather than policy-motivated, it does not matter whether they are workers or entrepreneurs. It would of course matter if we were to take the route of the so-called citizen-candidate models - see e. g. Osborne and Slivinski (1996) and Besley and Coate (1997).

¹⁷ We assume that the probability of voting is the same within each social group. The model could be extended to the case where the probability varies within each group for example because of an individual bias towards one of the candidates. If we adhere to the common assumption that biases are uniformly distributed, then the formal structure of our model continues to hold (see fn. 21).

the aggregate level—see for example Greene and Nikolaev (1999), Benabou (2000), and Horn (2011).¹⁸

Therefore, we assume that the probability of an individual participating in voting is positively related to her income. In our framework, this means that the entrepreneurs are more active than workers in the voting process, i.e. their abstention probability is lower. We take it that $\Gamma_w(\Delta V_i) < \Gamma_{\eta_k}(\Delta V_i)$ $k = y, m$, for any value of the difference in the utility from the two platforms.

Each candidate maximises (63) by choosing her own policy platform $\tau_{\eta}^j, \bar{e}^j, \bar{d}^j$ while taking the other candidate's platform as given; in a Nash equilibrium in which the candidates announce their policies simultaneously, the resulting equilibrium policies will be identical. As is well-known, then, the objective function of a candidate, that is (63), in a probabilistic voting model coincides with a generalised utilitarian social welfare function—see Mueller (2003).

In what follows, we will assume that for the income tax rate proposed by candidate j for entrepreneurs, τ_{η}^j , we always have interior solutions at the political equilibrium. In other words we assume that the abstention rate of the workers (who outnumber the entrepreneurs) is such as to guarantee an interior solution.¹⁹ As far as the education package is concerned, notice that there cannot exist an equilibrium in which both the workers and the entrepreneurs are constrained. If that were the case, one of the candidates could easily improve the outcome for both groups by simultaneously reducing the ration and the tax rates. For each candidate, the FOCs are:

$$[\tau_{\eta}^j] \quad \frac{\partial p^j}{\partial \tau_{\eta}^j} = \sigma \frac{\partial \Gamma_{\omega}}{\partial V_{\omega}} \frac{\partial V_{\omega}}{\partial \tau_{\eta}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \frac{\partial V_{\eta_y}}{\partial \tau_{\eta}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \frac{\partial V_{\eta_m}}{\partial \tau_{\eta}^j} = 0, \quad (65)$$

$$[\bar{e}^j] \quad \frac{\partial p^j}{\partial \bar{e}^j} = \sigma \frac{\partial \Gamma_{\omega}}{\partial V_{\omega}} \frac{\partial V_{\omega}}{\partial \bar{e}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \frac{\partial V_{\eta_y}}{\partial \bar{e}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \frac{\partial V_{\eta_m}}{\partial \bar{e}^j} \geq 0, \quad (66)$$

$$[\bar{d}^j] \quad \frac{\partial p^j}{\partial \bar{d}^j} = \sigma \frac{\partial \Gamma_{\omega}}{\partial V_{\omega}} \frac{\partial V_{\omega}}{\partial \bar{d}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \frac{\partial V_{\eta_y}}{\partial \bar{d}^j} + \frac{(1-\sigma)}{2} \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \frac{\partial V_{\eta_m}}{\partial \bar{d}^j} \geq 0, \quad (67)$$

where the derivatives of the indirect utility functions w.r.t. the policy parameters are given by (54)–(56) and (57)–(59).

Substituting for $\partial V_{\omega} / \partial \tau_{\eta}^j$ and $\partial V_{\eta_k} / \partial \tau_{\eta}^j$, $k = y, m$, from the preferred policies (54)–(56), condition (65) becomes:

¹⁸ Political economy models in line with this literature include Anderberg and Balestrino (2007), where the probability of abstention has been linked to the level of income, and Bourguignon and Verdier (2000), where the turnout in elections is determined by the level of education.

¹⁹ Corner solutions with $\tau_{\eta}^j = 1$, $\tau_{\omega}^j < 0$ are implausible in a democracy (as well as those with $\tau_{\eta}^j = 0$, $\tau_{\omega}^j > 0$).

$$\sigma \frac{\partial \Gamma_{\omega}}{\partial V_{\omega}} v' \frac{1-\sigma}{\sigma} - \frac{1-\sigma}{2} \lambda \left[\frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} (w_{\eta}^{1,t} + w_{\eta}^{2,t}) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} w_{\eta}^{2,t-1} \right] = 0, \quad (68)$$

$$\frac{\partial \Gamma_{\omega}}{\partial V_{\omega}} v' = \frac{\lambda}{2} \left[\frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} (w_{\eta}^{1,t} + w_{\eta}^{2,t}) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} w_{\eta}^{2,t-1} \right].$$

In words, the marginal tax rates are set so as to equalise the marginal utilities of income weighted by the sensitivity of the two citizen types' vote to the candidate's proposal at the equilibrium point,²⁰ that is when there is no difference in the proposed platforms (see Mueller 2003, ch. 12). Since the candidates' proposed policies are identical at the equilibrium, $\tau_{\eta}^A = \tau_{\eta}^B = \tau_{\eta}$, $\bar{e}^A = \bar{e}^B = \bar{e}$, and $\bar{d}^A = \bar{d}^B = \bar{d}$.

The intuition behind these results is the following. Let us suppose, for example, that $\partial \Gamma_{\eta_k} / \partial V_{\eta_k}$, $k = y, m$, is, for any given value of the difference between the utilities in (64), larger than $\partial \Gamma_{\omega} / \partial V_{\omega}$, meaning that entrepreneurs respond with a higher increase in the probability of voting for the candidate if the latter marginally differentiates her proposed platform in their favour; then, τ_{η} will be set in such a way that the marginal utility of income for the entrepreneurs, λ , is lower than the marginal utility of income for the workers, v' . That is, the policy favours the citizen whom the candidate perceives as more likely to vote for her as a consequence of such a favour.²¹

In (68) we consider interior solutions for τ_{η} . Notice, however, that the characteristics of the solution depend on the abstention rate of the workers. Considering that there are more workers than entrepreneurs, in general τ_{ω} cannot be positive unless the workers' abstention rate is particularly high even for large differences in the utility they can obtain from the two candidates' platforms,

$$V_i \left(\tau_{\eta}^j, \tau_{\omega}^j \left(\tau_{\eta}^j, \bar{e}^j, \bar{d}^j \right), \bar{e}^j, \bar{d}^j \right) - V_i \left(\tau_{\eta}^{-j}, \tau_{\omega}^{-j} \left(\tau_{\eta}^{-j}, \bar{e}^{-j}, \bar{d}^{-j} \right), \bar{e}^{-j}, \bar{d}^{-j} \right), \quad i = \omega, \eta_y, \eta_m \quad (69)$$

The educational policy must usually be paired with a redistributive taxation in favour of the workers because the latter suffer from a reduction in the household public good. This scenario can arise if the entrepreneurs' benefits coming from the workers' education are sufficiently high to compensate both the cost of the compulsory education package and the redistributive policy. Clearly, such a cost would be lower if the workers attached positive value to their children's education. In that case, less redistribution would be needed for the workers to accept the compulsory

²⁰ Notice, however, that for the entrepreneurs the term in square bracket is the average of the sensitivity of the young and the mature ones.

²¹ In the case where the voting probability varies within each social group because of a uniformly distributed individual bias towards one of the candidates, the sensitivity $\partial \Gamma_i / \partial V_i$ represents the height of the distribution. Such height will be higher, the lower is the range over which the bias is distributed. Consequently, the sensitivity is higher the more concentrated is the distribution, implying that the policy will favour the group whose probability of voting is less dispersed (see Mueller 2003, ch. 12).

education policy. If education were highly valued by the workers, the equilibrium policy could even prescribe positive values for τ_ω .

When workers do not attach any value to their children's education, substituting (55), (58), and (61) into condition (66) and (56), (59), and (62) into condition (67), the two conditions can be re-written as

$$[\bar{e}] \quad -\frac{\partial \Gamma_\omega}{\partial V_\omega} v' + \frac{1-\sigma}{2} \left\{ \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \left[(1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{e}} + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{e}} \right] + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{e}} \right\} \geq 0, \quad (70)$$

$$[\bar{d}] \quad -\sigma \frac{\partial \Gamma_\omega}{\partial V_\omega} f' + \frac{1-\sigma}{2} \left\{ \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \left[(1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{d}} \lambda + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{d}} \right] + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{d}} \right\} \geq 0. \quad (71)$$

Further, substituting (65), the above equations can be re-written as

$$[\bar{e}] \quad \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} (1-\sigma) \left((1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{e}} + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{e}} \right) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} (1-\sigma) \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{e}} \geq \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \lambda (w_\eta^{1,t} + w_\eta^{2,t}) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \lambda w_\eta^{2,t-1}; \quad (72)$$

$$[\bar{d}] \quad \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} (1-\sigma) \left((1-\tau_\eta) \frac{\partial w_\eta^{2,t}}{\partial \bar{d}} \lambda + \beta \frac{\partial w_\eta^{3,t}}{\partial \bar{d}} \right) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} (1-\sigma) \beta \frac{\partial w_\eta^{3,t-1}}{\partial \bar{d}} \geq \frac{\partial \Gamma_{\eta_y}}{\partial V_{\eta_y}} \sigma \lambda \frac{f'}{v'} (w_\eta^{1,t} + w_\eta^{2,t}) + \frac{\partial \Gamma_{\eta_m}}{\partial V_{\eta_m}} \sigma \lambda \frac{f'}{v'} w_\eta^{2,t-1}. \quad (73)$$

On the l.h.s. of (72) we have a measure of the marginal benefit of educational expenditure, for the young and the mature entrepreneurs weighted by the respective vote sensitivities to the candidates' proposals; on the RHS, we have the weighted marginal cost, expressed in utility terms for $\bar{e}$. Similarly on the l.h.s. of (73) we have a measure of the weighted marginal benefit of school time while on the RHS we have the weighted marginal cost expressed as an opportunity cost for $\bar{d}$.

Then, as long as marginal benefits exceed or equal marginal costs, a solution in which a certain level of compulsory education is enforced emerges. We have then the interesting result that a compulsory education policy may be implemented at the political equilibrium, despite the fact that one of the two groups of which the society is composed would not educate the children in a free-market equilibrium. The driving force behind this result is the fact that the entrepreneurs gain from having an educated workforce.

The reputational effect plays a role in this respect, by making mature entrepreneurs care about future profits and thus value workers' education, but it

is by no means necessary to achieve our general result. In the absence of such effect (i.e. for $\beta = 0$), the l.h.s. of (72) and (73) would be lower yielding lower equilibrium levels for $\bar{e}$ and $\bar{d}$. The benefit of the policy would be enjoyed only by young entrepreneurs implying a milder policy in case the equilibrium exists and a lower probability of existence of an equilibrium with the above characteristics but the general conclusion would be unaffected.

5 Conclusions

Over the years, there have been several contributions to the political economy of education. However, their focus seems to have been mostly on secondary or tertiary education. Also, typically, the main driving force behind the results has been the presence of income dispersion. Consider for example, the work by Epple and Romano (1996b). In their model, a publicly provided private good, which could be education, is financed through a flat-rate income tax and policy is determined by majority rule; agents differ by their fixed incomes. At the political equilibrium, the private good is publicly provided as long as it is possible to top it up; interestingly, for some preference configurations, the political equilibrium is of the “ends-against-the-middle” variety, i.e. low- and high-income agents favour low levels of public provision whereas the middle-income agents favour high levels of public provision (see also Epple and Romano 1996a). Another well-established result is that post-compulsory education policies are at least partially regressive, redistributing income from the lower income groups to middle- and high income groups (see e.g. Fernandez and Rogerson 1995).

We have taken a different route here, paying attention specifically to the question whether education should be made compulsory or not. We considered an economy with two categories of agents: entrepreneurs and workers. The type of occupation, rather than the income dispersion, plays a crucial role in the analysis. In *laissez-faire*, the former gain from having their children educated, while the latter have no interest in sending their children to school. We characterised the preferred education policy-cum-redistributive taxation for the two groups, and find that entrepreneurs favour a compulsory education policy while workers prefer a purely redistributive taxation scheme (in both cases, the policy should preferably be financed entirely by the other group). Then, we introduced a political process with probabilistic voting and verified that an equilibrium with both a compulsory education policy and some redistribution may exist in which the workers are constrained but the entrepreneurs, who benefit from hiring educated workers, are not.

We should note that, on top of establishing a political economy rationale for compulsory schooling based on the intuition that due to a productivity reason entrepreneurs care for education more than workers, our model shows also that it must co-exist with a (limited) redistribution policy. In particular, the model allows to make the point that a certain amount of redistribution is needed to compensate the workers for the fact that they are forced to “over-educate” their kids. This kind of perspective on redistribution could only be achieved, we believe, within the structure of the present model.

To the best of our knowledge, the literature on this topic, at least if we consider the political economy line of work, is limited. Most papers follow different approaches from ours. As an example of these alternative views, consider the contribution by Gradstein (2000), whose elegant argument is based on the idea of time inconsistency. He argues that private financing of education can be an inferior public choice if the current government representing the parents is unable to pre-commit the next generation to a restrained redistributive policy. He observes that public education, relative to private education, generates a more equal income distribution for the children, and therefore suggests that in the future the government will have to implement a relatively moderate redistributive policy, as chosen by the median voter. This reduces the incentive to under-invest in the children's education, incentive that instead would be large in case the parents expected a more aggressive redistribution policy. Thus, human capital should be accumulated at a faster pace under a public education regime, and this would make it preferable for a majority of parents to the alternative of a private education regime. Another interesting example of a paper in the same vein is Correa et al. (2020), where the political support for different education funding regimes in a one-person, one-vote system is studied. Free education, in which all pupils are treated equally (the same amount of resources is spent on each of them) turns out to be the Condorcet winner. This is because, in economies with some degree of income inequality, any other system concentrates the educational expenditure in some way, either favouring the richest families or those with the smartest kids, and therefore, lacks majority support. This provides a political economy explanation for the observation that governments tend to favor free education for all students (i.e., to spend the same amount on each student).

Clearly, ours is an entirely different line of reasoning, not in contradiction, but certainly based on other foundations and moreover focused specifically on whether education should be made compulsory or not, rather than on whether it should be financed by the State or not (which is of course a somewhat different issue). In our model, income disparity plays a part, in particular by supporting the assumption that entrepreneurs are not constrained by compulsory education, but what really drives our result on the desirability of compulsory education at the political equilibrium is the difference in occupation, i.e. the different role that education plays for the entrepreneurs as opposed to the workers because of the industrialization/productivity reason we have mentioned in the Introduction. Redistribution, however, by compensating workers for the loss of their children's production, plays a role in that it makes workers accept the presence of compulsory education at the political equilibrium.

Funding Open access funding provided by Università di Pisa within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended

use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderberg D, Balestrino A (2007) Non-cooperative households and the size and composition of public expenditure. *Econ Gov* 8:61–81
- Balestrino A, Grazzini L, Luporini A (2017) A normative justification of compulsory education. *J Popul Econ* 30:537–567
- Bandiera O, Mohnen M, Rasul I, Viarengo M (2018) Nation-building through compulsory schooling during the age of mass migration. *Econ J* 128:1–48
- Benabou R (2000) Unequal societies: income distribution and the social contract. *Am Econ Rev* 90:96–129
- Becker S, Woessmann L (2010) The effect of protestantism on education before the industrialization: evidence from 1816 Prussia. *Econ Lett* 107(2):224–228
- Besley T, Coate S (1997) An economic model of representative democracy. *Q J Econ* 112:85–114
- Bourguignon F, Verdier T (2000) Oligarchy, democracy, inequality and growth. *J Dev Econ* 62:285–313
- Bovenberg AL, Jacobs B (2005) Redistribution and education subsidies are Siamese twins. *J Pub Econ* 89:2005–2035
- Checchi D (2005) The economics of education. Human capital, family background and inequality. Cambridge University Press, Cambridge
- Cigno A (2013) How to deal with covert child labor and give children an effective education, in a poor developing country. *World Bank Econ Rev* 26:61–67
- Cigno A, Giannelli G, Rosati F (1998) Voluntary transfers among Italian households: altruistic and non-altruistic explanations. *Struct Chang Econ Dyn* 4:435–53
- Cigno A, Giannelli G, Rosati F, Vuri D (2006) Is there such a thing as a family constitution? A test based on credit rationing. *Rev Econ Househ* 4:183–204
- Cipolla CM (1969) Literacy and development in the west. Penguin, Harmondsworth
- Ciuffetti A (2004) Casa e lavoro (in Italian). Giada, Narni
- Correa JA, Lu Y, Parro F, Villena M (2020) Why is free education so popular? A political economy explanation. *J Pub Econ Theory* 22:973–991
- Dee T (2004) Are there civic returns to education? *J Pub Econ* 89:1697–1720
- Dellheim C (1987) The creation of a company culture: Cadburys, 1861–1931. *Am Hist Rev* 92(1):13–44
- Epple D, Romano R (1996a) Ends against the middle: determining public service provision when there are private alternatives. *J Pub Econ* 62:297–325
- Epple D, Romano R (1996b) Public provision of private goods. *J Polit Econ* 104:57–84
- Fyfe A (2005) Compulsory education and child labour: historical lessons. Contemporary challenges and future directions, ILO Working Draft
- Fellman S (2019) Company welfare programmes and industrial modernization in Finland. *Scand J Hist* 44(5):665–692
- Fernandez R, Rogerson R (1995) On the political economy of education subsidies. *Rev Econ Stud* 62:249–262
- Galor O (2006) From stagnation to growth: a unified growth theory. In: Aghion P, Durlauf S (eds) *Handbook of economic growth*. North Holland, Amsterdam
- Galor O, Moav O (2006) Das human-kapital: a theory of the demise of the class structure. *Rev Econ Stud* 73:85–117
- Galor O, Moav O, Vollrath D (2009) Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *Rev Econ Stud* 76:143–179
- Gradstein M (2000) An economic rationale for public education: the value of commitment. *J Monet Econ* 45:463–474
- Greene K, Nikolaev O (1999) Voter participation and the redistributive state. *Pub Choice* 98:213–226
- Horn D (2011) Income inequality and voter turnout, GINI Discussion Paper no. 16
- Juif D (2019) Mining, paternalism and the spread of education in the Congo since 1920, Universidad Carlos III de Madrid, African Economic History Working Paper series no. 46/2019

- Lindbeck A, Weibull J (1987) Balanced-budget redistribution as the outcome of political competition. *Pub Choice* 52:273–297
- Milligan K, Moretti E, Oreopoulos P (2004) Does education improve citizenship? Evidence from the United States and the United Kingdom. *J Pub Econ* 88:1667–95
- Mueller D (2003) *Public choice III*. Cambridge University Press, Cambridge
- Murtin F, Viarengo M (2011) The expansion and convergence of compulsory schooling in Western Europe, 1950–2000. *Economica* 78:501–522
- Naito K, Nishida K (2017) Multistage public education, voting and income distribution. *J Econ* 120:65–78
- Osborne MJ, Slivinski A (1996) A model of political competition with citizen-candidates. *Q J Econ* 111:65–96
- Reid D (1985) Industrial paternalism: discourse and practice in nineteenth-century French mining and metallurgy. *Comp Stud Soc His* 27(4):579–607
- Squicciarini MP (2020) Devotion and development: religiosity, education, and economic progress in nineteenth-century France. *Am Econ Rev* 110(11):3454–3491
- Tone A (1997) *The business of benevolence: industrial paternalism in progressive America*. Cornell University Press, Ithaca
- Tsutsui W (1997) Rethinking the paternalist paradigm in Japanese industrial management. *Bus Econ Hist* 25(2):561–572

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.