The Political Economy of Intergenerational Income Mobility *

Andrea Ichino
UNIVERSITY OF BOLOGNA

Loukas Karabarbounis
HARVARD UNIVERSITY

Enrico Moretti
UNIVERSITY OF CALIFORNIA, BERKELEY

December 2009

Abstract

The intergenerational elasticity of income is considered one of the best measures of the degree to which a society gives equal opportunity to its members. While much research has been devoted to measuring this reduced-form parameter, less is known about its underlying structural determinants. Using a model with exogenous talent endowments, endogenous parental investment in children and endogenous redistributive institutions, we identify the structural parameters that govern the intergenerational elasticity of income. The model clarifies how the interaction between private and collective decisions determines the equilibrium level of social mobility. Two societies with similar economic and biological fundamentals may have vastly different degrees of intergenerational mobility depending on their political institutions. We offer empirical evidence in line with the predictions of the model. We conclude that international comparisons of intergenerational elasticity of income are not particularly informative about fairness without taking into account differences in politico-economic institutions.

JEL-Code: E24, J62, J68, P16.
Keywords: Intergenerational Mobility, Public Education, Political Institutions.

*We would like to thank Alberto Alesina, Giacomo Calzolari, Ed Glaeser, John Hassler, Larry Katz, Mattia Nardotto, Sevi Rodriguez-Mora and seminar participants at C6-Capri, ESSLE-CEPR 2008, Edinburgh University, Harvard and Sorbonne for useful discussions. We also thank two anonymous referees for their helpful suggestions.

Email: andrea.ichino@unibo.it; karabarb@fas.harvard.edu; moretti@econ.berkeley.edu.

1 Introduction

The intergenerational elasticity of income is generally considered one of the best summary measures of the degree to which a society gives equal opportunities of success to all its members, irrespective of their family background. Starting with pioneering work by Solon (1992) and Zimmerman (1992), the economic literature has made important advances on the question of how to *measure* this parameter using the Galton-Becker-Solon regression:

$$y_s = a + \beta y_f + u_s \tag{1}$$

where y_s is son's log income and y_f is father's log income. A lower β denotes a smaller association between father's and son's income and therefore a higher degree of social mobility. As such, a lower β is often interpreted as being a desirable feature of a society.

While we have learned a lot on how to estimate this reduced-form parameter, less progress has been made on understanding its underlying structural determinants. What does β actually measure? Is a lower β necessarily more desirable? Important progress on these questions has been made by Becker and Tomes (1979), who have shown how the intergenerational persistence of income reflects both "nature and nurture". In their model individuals are assigned talent by nature, and parents can add to that talent by privately investing in their children. The intergenerational transmission of income is therefore a combination of exogenous biological factors and endogenous optimizing behavior of parents. However, the Becker and Tomes model generally ignores the role of redistributive policies and their deeper determinants. Redistributive policies have the potential to play an important role in determining how income is transmitted from one generation to the next. For example, public education can significantly affect economic opportunities of individuals who come from disadvantaged socioeconomic backgrounds. At the same time, it can also affect parents' incentives to privately invest in their children human capital, both directly and throughout the disincentive effect of taxes. More in general, most redistributive policies—including taxation, affirmative action, welfare programs, subsidies that target poor individuals—potentially affect the intergenerational elasticity of income. While some studies have highlighted the role of public policies as a determinant of social mobility, most existing studies take these policies as *exogenous*.

In this paper we use a model with exogenous talent endowments, endogenous parental investment in children and endogenous redistributive institutions, to identify the structural parameters that govern the intergenerational mobility. Our framework extends the Becker and

Tomes framework and clarifies how the interaction between private and collective decisions determines the equilibrium level of social mobility. The model allows for a structural interpretation of the widely studied parameter β. This is important because it allows a better understanding of the deeper politico-economic determinants of intergenerational mobility and the role of public policy. The model also shows how we should interpret and rank differences over time and across countries in β. Since redistributive policies generate a trade-off between insurance and incentives, the optimal β is not necessarily zero for all societies. In addition, international comparisons of intergenerational elasticity of income are shown to be not particularly informative about fairness without taking into account differences in politico-economic institutions. The predictions of the model seem generally consistent with the empirical evidence.

Our framework focuses on how parents transfer economic endowments to their children through private and collective investment in their human capital. Before having children, parents know their own genetic ability but are uncertain about their children genetic ability. Consistent with Becker and Tomes (1979) and Loury (1981), parents can decide to invest privately in the human capital of their children, given an exogenous degree of transmission of genetic ability. This private investment offsets some of the risk of having low genetic ability, thus reducing the probability that an individual might turn out to have low productivity and therefore low income. Since private investment can offset some but not all of the genetic risk, parents "under the veil of ignorance" have an incentive to collectively create public institutions that provide further insurance against the risk of low genetic ability. A natural example of this type of policy is public education.

We model public education as an insurance system that increases the income of the low talented children, at the expense of lowering the income of the more talented children. We show how and why a more progressive educational policy increases social mobility in equilibrium. The equilibrium level of social mobility depends on the costs and benefits of public education. This trade-off is resolved by two forces: (i) the balance between costly insurance and incentives to privately invest in children's human capital and (ii) the political process that aggregates conflicting interests regarding the desired degree of social mobility.

A novel insight of our analysis is to show how political economy forces shape the equilibrium level of social mobility. Even if public education is relatively costless to provide for the average family, it may hurt the interests of the rich dynasties who, in a world of increased social mobility, are more likely to move down the income ladder. As a result, the maximum amount of mobility ($\beta = 0$) is not necessarily the equilibrium one, even when public insurance is relatively cheap to provide.

More generally, the model shows that existing differences in β across countries are (at least in part) governed by all those political institutions that affect public education. Therefore, two societies with similar fundamentals (such as the degree of parental altruism, variability in market earnings, degree of biological and cultural transmission of family characteristics, labor market discrimination, asset market incompleteness etc.) may display very different degrees of intergenerational mobility depending on the identity of the politically decisive family.

In the last part of the paper, we use data on a cross section of countries for which reliable estimates of β are available to test the predictions of the model. In general, we find that they are consistent with the empirical evidence. For example, our model predicts that in countries where rich dynasties are more politically active than poor dynasties, social spending for public education should be lower and therefore income mobility should be lower. We find that this appear to be the case in our sample. The difference in the probability of party affiliation between rich and poor appears to be strongly correlated with β. Such difference has five times larger predictive power than the rate of return to education, which is often considered as one of the most prominent determinants of mobility (Solon, 1999, 2004; Corak, 2006). While causality is obviously unclear, these empirical correlations are at least consistent with our model.

The rest of the paper is organized as follows. Section 2 discusses the related literature. In Section 3 we describe the model and examine its positive properties. In Section 4 we derive the politico-economic determinants of social mobility and show their relation to the Galton-Becker-Solon regression. In Section 5 we present our empirical evidence. Section 6 concludes. All omitted derivations are in Appendix 1. Appendix 2 describes the data.

2 Related Literature

The objective of our model is to derive the structural politico-economic parameters underlying the intergenerational elasticity of income. This coefficient—β in equation (1)—has been the main focus of the existing empirical literature, see among others Solon (1992), Zimmerman (1992), Björklund and Jäntti (1997), Mulligan (1997) and Solon (1999). Our model is also related to a more recent empirical strand of research that examines within-country trends in mobility and compares β over time, see for instance Mazumder (2005, 2007), Lee and Solon (2006), and Aaronson and Mazumder (2008).

Most theoretical work in this area has focused on the role of the genetic transmission of ability, the incentives for parental investment, and the role of the asset market in explaining the intergenerational transmission of income. Our framework builds on the theoretical work of Becker and Tomes (1979), and on extensions of this work by Goldberger (1989), Mulligan (1997) and Solon (2004).

While some studies have highlighted the role of public policies as a determinant of social mobility, most existing studies take these policies as exogenous. Examples of papers that have argued that institutions may be important determinants of mobility, but take these institutions as exogenous include, among others, the original contribution of Becker and Tomes (1979), Glomm and Ravikumar (1992), Checchi, Ichino and Rustichini (1999), Solon (1999, 2004), Davies, Zhang and Zeng (2005), Mayer and Lopoo (2005), and Hassler, Rodriguez Mora, and Zeira (2007).

In our setting, social mobility depends on public redistributive policies that we model as the outcome of a politico-economic equilibrium. In this sense, our model relates to the equilibrium models of Saint-Paul and Verdier (1993), Alesina and Rodrik (1994), and Persson and Tabellini (1994). These papers show how cross sectional inequality causes growth, through endogenous public policies. Benabou (1996) further develops this strand of literature and endogenizes the relationship between inequality, social mobility, redistribution and growth as a function of the incompleteness of the financial market. While our model abstracts from (physical) capital accumulation, it emphasizes the endogenous production of human capital (talent) as an intermediate input for the production of final income. Fernandez and Rogerson (1998) analyze a reform from a locally financed to a centralized educational system in a multicommunity model with endogenous choice of location. Relative to their paper, we instead focus on explaining cross country outcomes. In this case, migration becomes a less important determinant of social mobility and differences in political institutions become stronger determinants of social mobility. As in our paper, Bernasconi and Profeta (2007) endogenize institutions in a model with mobility and argue that the politically-determined level of public education may reveal the true talent of the children and relax the mismatch of talents to occupations. Relative to this paper, our model includes both economic and political choices.

In a seminal paper, Piketty (1995) explains the emergence of permanent differences in attitudes toward redistribution. Benabou and Ok (2001) show how rational beliefs about one's relative position in the income ladder affect the equilibrium level of redistribution. These papers derive the implications of social mobility for redistributive policies, while we focus on the reverse channel. Specifically, we analyze how endogenously chosen public policies affect the intergenerational mobility.

It is important to note that because the direction of causation in our model differs from the one emphasized in Benabou and Ok (2001), we obtain a different prediction for the relationship between mobility and redistribution in the US and Europe. In their paper, more mobility is associated with less redistribution because voters who are below the mean oppose redistribution in the rational expectation of income gains in the future. This explanation is intuitive, but cross Atlantic evidence suggests that the US is less mobile and less redistributive than continental Europe.[1] In our paper, political economy forces that constraint the development of public education also lead to a lower degree of social mobility. Thus, our model predicts a positive correlation between social mobility and redistribution of income across countries.

3 A Simple Model of the Intergenerational Transmission of Income

We first setup the model and derive the intergenerational transmission equation for income and talent. Then, we derive the first and second moments of income and talent distributions and discuss how these moments evolve in response to more progressive public policies.

3.1 Set-up of the Model

We consider an infinite horizon overlapping generations economy populated by a measure one of dynasties, $i \in [0, 1]$. In each period $t = 0, 1, 2, ...$ two generations are alive, fathers and sons. In each generation, earnings (which we also call "output" or "income" interchangeably) are produced according to the production function $Y_{i,t} = f(\mu_t, \Theta_{i,t}, U_{i,t})$. The parameter μ_t represents the public policy; $\Theta_{i,t}$ is father's human capital or basic skill (e.g. IQ) which we call "talent"; and $U_{i,t}$ denotes a random and inelastic factor of production which represents "market luck". Specifically, we assume that the production function is given by:

$$Y_{i,t} = \mu_t^{\alpha} \left(U_{i,t}\Theta_{i,t}\right)^{\mu_t} \tag{2}$$

where $\mu_t \in (0, 1]$ and $\alpha \geq 0$.

Figure 1 shows the production function graphically. Public policy and its effects are characterized by two parameters, μ_t and α. The parameter μ_t characterizes the amount of redistribution in the economy. A lower μ_t implies a more progressive public policy, but also more

[1] See the evidence in Section 5. See also Alesina and Glaeser (2004) for more on this point.

distortions. This is shown visually in the left panel of Figure 1, where for given amount of talent and market luck, a lower μ_t is associated with less output for the talented or lucky families, but with more output for the less talented or unlucky families. The most natural example of the public policy represented by μ_t is public education. In Section 5 we offer evidence in line with this interpretation of μ_t.[2] Henceforth, a *lower* μ_t is called a more progressive public policy or a more progressive educational system.

The parameter α characterizes the efficiency of public education. For a given μ_t, a higher α implies that a smaller fraction of talents $\Theta_{i,t}$ gains from progressivity because the system creates disincentives for high talented agents. In the right panel of Figure 1, the area to the left of the intersection of the production function with the 45 degree line measures the gains from progressivity. As α increases, this area becomes smaller relative to the area to the right of the intersection of the production function with the diagonal, which measures the efficiency costs of progressivity.[3] Henceforth, a higher α denotes more distortions.

In each period t the following events take place:

1. Fathers produce output $Y_{i,t}$ according to equation (2), given the predetermined talent $\Theta_{i,t}$, market luck $U_{i,t}$ and public policy μ_t.

2. Fathers choose the policy for their sons, μ_{t+1}, according to the institution or political process P.

3. Sons are born with a random family endowment $V_{i,t+1}$. The random factor of production $U_{i,t+1}$ is realized.

4. Fathers observe $V_{i,t+1}$ and $U_{i,t+1}$ and choose investment $I_{i,t}$ to maximize the dynastic utility, given resources $Y_{i,t}$. Investment produces son's talent according to the production function $\Theta_{i,t+1} = g(I_t, h_i V_{i,t+1})$.

5. Fathers die, sons become fathers and the process repeats ad infinitum.

For this Section we treat μ as an exogenous parameter. In Section 4 we endogenize it. Son i is born with random family endowment $V_{i,t+1}$, which, following Becker and Tomes (1979), is

[2] Pekkarinen, Uusitalo and Kerr (2008) show how the major Finnish educational reform in the 1970s decreased the intergenerational elasticity of income from 0.30 to 0.23. Their finding is consistent with our interpretation of μ_t.

[3] We do not restrict $\Theta_{i,t}$ to be smaller than unity. If in some period $\Theta_{i,t} \leq 1$ for all families i, we can think the special case with $\alpha = 0$ as a growth-enhancing reform that benefits every family, with the least talented families gaining relatively more.

assumed to follow a "Galtonian" AR(1) process:

$$v_{i,t+1} = (1 - \rho_1)\rho_0 + \rho_1 v_{i,t} + \epsilon_{i,t+1} \tag{3}$$

where $v = \ln V$ (small caps denote logs of corresponding variables throughout the paper). For every dynasty i, $\epsilon_{i,t+1}$ is a white noise process with expected value $\mathbf{E}(\epsilon_{i,t}) = 0$, variance $\mathbf{Var}(\epsilon_{i,t}) = \sigma_v^2$ and zero autocorrelations. We have $0 \leq \rho_1 < 1$ and therefore the logarithm of family endowment regresses towards the mean, has stationary expectation $\mathbf{E}(v_{i,t}) = \rho_0$, and has stationary variance $\mathbf{Var}(v_{i,t}) = \sigma_v^2/(1 - \rho_1^2)$. The parameter ρ_1 characterizes the cultural or genetic inheritance of traits related to talent and income, and is assumed identical across families i.

A second random component is represented by market luck, $U_{i,t+1}$, whose logarithm is a white noise process, has variance σ_u^2, and is independent to $\epsilon_{i,t}$. The difference between $U_{i,t}$ and $\Theta_{i,t}$ is that the latter is an elastic factor of production. As a result, talent is affected by the inefficiencies associated with the policy μ.

Fathers care about the quality of their children. They observe $V_{i,t+1}$ and $U_{i,t+1}$ and decide how to allocate their predetermined income $Y_{i,t}$ into consumption $C_{i,t}$ and investment $I_{i,t}$ in order to maximize the dynastic utility:

$$\ln C_{i,t} + \frac{1}{\gamma} \ln Y_{i,t+1} \tag{4}$$

subject to the budget constraint:

$$C_{i,t} + I_{i,t} = Y_{i,t} \tag{5}$$

where $Y_{i,t+1}$ is children's income.[4] The parameter $\gamma > 0$ captures the degree of parental altruism, with higher values denoting smaller altruism. Parental investment $I_{i,t}$ can be thought as an private educational input (e.g. tuition fees) that increases a child's talent.

Sons' talent is produced with the following production function:

$$\Theta_{i,t+1} = (h_i V_{i,t+1}) I_{i,t} \tag{6}$$

where h_i is a family-specific time-invariant ability effect which allows dynasties to be *ex-ante* heterogeneous. This heterogeneity captures long-run differences in market incomes, for instance due to labor market discrimination against certain racial, ethnic or religious groups. We assume that h_i is distributed according to the density function ϕ_h with bounded support $\mathbf{H} \subset R_{++}$, and is orthogonal to the disturbances $\epsilon_{i,t+1}$ and $u_{i,t+1}$.

[4] We assume that fathers cannot borrow against their son's future income. See Loury (1981), Becker and Tomes (1986) and Mulligan (1997), for an analysis of the relationship between social mobility and borrowing constraints. See also Benabou (1996, 2000).

3.2 The Transmission of Income Across Generations

In this Section we restrict attention to steady state public policies, i.e. we set $\mu_{t+1} = \mu_t = \mu$ for all t. Under this assumption, income and talent are stochastic processes with well defined and easy to analyze unconditional stationary moments. We generalize our analysis in Section 4, where we endogenize the choice of μ. Solving the problem in (4)-(5), using the production functions (2) and (6), and taking logs, we obtain the equation that describes the intergenerational transmission of income in family i:

$$y_{i,t+1} = \delta_{0,i} + \delta_1 y_{i,t} + \delta_2 v_{i,t+1} + \delta_3 u_{i,t+1} \tag{7}$$

where:

$$\delta_{0,i} = \delta_0 + \delta_i \tag{8}$$

$$\delta_0 = \mu \ln \left(\frac{\mu}{\mu + \gamma} \right) + \alpha \ln \mu \tag{9}$$

$$\delta_i = \mu \ln h_i \tag{10}$$

$$\delta_1 = \mu \tag{11}$$

$$\delta_2 = \mu \tag{12}$$

$$\delta_3 = \mu \tag{13}$$

The intercept $\delta_{0,i}$ can be decomposed into two parts. δ_0 is a common effect across all dynasties i, and δ_i is the dynasty-specific time-invariant effect due to h_i. Our autoregressive coefficient, δ_1, is different from the one in Becker and Tomes (1979) because we assume multiplicative (in levels) production functions for output and talent.[5] While the previous literature has focused on the role of private incentives for the intergenerational mechanism, our δ_1 coefficient emphasizes instead the role of public policies. Specifically, the novel element of our model

[5] Goldberger (1989) explains in detail the difference between the additive production function (as in the Becker and Tomes model) and the multiplicative production function. We also note that in our specific Cobb-Douglas environment, the degree of parental altruism (γ) does not enter into the intergenerational transmission equation directly, i.e. for given policy μ (see Solon, 2004, for a similar result).

is that the slope δ_1 is *collectively* decided by the fathers of each dynasty. Therefore, our mechanism maps collective action outcomes to equilibrium levels of intergenerational transmission of income. In the Appendix we present the intergenerational transmission of talent.

3.3 The Trade-Off Between Equity and Efficiency

3.3.1 Expectations

From (7) we take the unconditional, stationary expectation of income ("long-run income") for family i:

$$\mathbf{E}(y_{i,t+1}|h_i) = \frac{\mu\left[\rho_0 + \ln h_i + \ln\left(\frac{\mu}{\mu+\gamma}\right)\right] + \alpha \ln \mu}{1-\mu} \tag{14}$$

for all t. In (14), the expectation is conditioned only on h_i to denote the dependency of long-run income on long-run family ability h_i. There are four ways through which the public policy μ affects long-run income.

1. *Distortions in Private Investment*: This is captured by the $\ln\left(\frac{\mu}{\mu+\gamma}\right)$ term. When public policy becomes more progressive (lower μ), the marginal propensity to invest in human capital, $\mu/(\mu+\gamma)$, is lower and as a result the long-run level of income tends to decline. This effect is identical for every dynasty i.

2. *Direct Distortions in Output*: This effect is shown in the $\alpha \ln \mu$ term, and is associated with the shifter μ^α in the production function for income in equation (2). The effect of μ on output is more adverse when the parameter α increases.

3. *Social Insurance or Benefits of Public Education*: The μ term that multiplies the bracket in the numerator of (14) captures the exponent of the term Θ^μ in equation (2). For low ability dynasties (low h_i), a more progressive public educational system increases long-run income. The opposite happens for sufficiently high ability families. The intuition is shown in Figure 1.

4. *Intertemporal Insurance or Social Mobility*: This effect is given by the denominator $1-\mu$ and is associated with the slope of the intergenerational transmission of income δ_2 in (7). For sufficiently low ability dynasties (low h_i), the numerator is negative and the prospect of upward mobility (lower μ) increases long-run income. For high ability dynasties, the numerator is positive and increased mobility decreases their long-run income.

We can write father i's conditional (on the state of the system) expectation for son's income as the sum of the long-run level of income in (14) and the transitory deviation of current income and current family endowment from their long-run levels:

$$\mathbf{E}_t(y_{i,t+1}|h_i) = \mathbf{E}(y_{i,t+1}|h_i) + \mu\,(y_{i,t} - \mathbf{E}(y_{i,t+1}|h_i)) + \mu\rho_1\,(v_{i,t} - \rho_0) \tag{15}$$

where the time subscript in the left hand side denotes conditioning on the information set as of period t (which is summarized by father's income, $y_{i,t}$, and family endowment, $v_{i,t}$). As we show in Section 4.1, fathers take into account how progressivity affects this conditional expectation when voting for μ.

This analysis highlights two important points. First, there is a trade-off between equity and efficiency. Second, there is political conflict over the equilibrium level of social mobility. In particular, as we discuss more formally in Section 4.2, fathers with higher ability h_i or with favorable shocks in their market activity, $u_{i,t}$, or in their family endowment, $v_{i,t}$, prefer less progressive policies. It is this heterogeneous effect of μ on dynastic welfare that makes the political economy aspect of the model interesting and supports our argument that politico-economic determinants may be significantly associated with mobility outcomes.

3.3.2 Variances

To understand the implication of our model for inequality, we first consider the stationary, unconditional variability that a given dynasty h_i faces in its income process. From equation (7) this is:

$$\mathbf{Var}(y_{i,t+1}|h_i) = \frac{\mu^2}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 \tag{16}$$

Inequality across generations occurs because the disturbances $\epsilon_{i,t+1}$ and $u_{i,t+1}$ have different realizations across time for a given family i. From inspection of (16), we see that a more progressive system (lower μ) reduces the variability of income. In addition, it lowers the fraction of variability attributed to family luck $v_{i,t+1}$. Intuitively, market luck $u_{i,t+1}$ matters only for the final production of income, while family luck $v_{i,t+1}$ affects both the production of talent directly, and the production of final output indirectly (through talent). As a result, more progressive public policies reduce the relative importance of the latter in the intergenerational variance of income.

If all families were identical, then the variance that families face across generations in (16) coincides with the stationary inequality in the cross section of families. More in general, with heterogeneous families, the *ex-post* or cross-sectional variance of income can be decomposed in two parts:[6]

$$\mathbf{Var}(y_{i,t+1}) = \mathbf{Var}(y_{i,t+1}|h_i) + \mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i)) \tag{17}$$

The second term in (17) represents the variance "under the veil of ignorance", which from (14) equals:

$$\mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i)) = \frac{\mu^2 \mathbf{Var}(\ln h_i)}{(1-\mu)^2} \tag{18}$$

To summarize, in (17) the stationary total inequality in the cross section of families is decomposed into the dynastic variability in the process for income—common to all families i—and the inequality that arises because heterogeneous families have different levels of long-run income. It is immediate to see that a more progressive educational system reduces all inequalities. In the Appendix we also discuss the variance of talent.

3.3.3 Covariances

Consider now the intergenerational correlation of income. This summary statistic is what the literature calls social mobility, inequality across generations or "equality of opportunity". Conditioning on h_i, we distinguish the intergenerational correlation of income within family, $\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)$, from the correlation we may observe in the data when families are heterogeneous, $\mathbf{Corr}(y_{i,t+1}, y_{i,t})$. The latter is discussed in Section 4.3 in relation to the Galton-Becker-Solon regression. Consider the time series of output and talent for some family i with time-invariant ability level h_i. Given that we are in a stationary state with $\mathbf{Var}(y_{i,t+1}|h_i) = \mathbf{Var}(y_{i,t}|h_i)$, we can derive the dynastic intergenerational correlation of income:

$$\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i) = \frac{\mathbf{Cov}(y_{i,t+1}, y_{i,t}|h_i)}{\mathbf{Var}(y_{i,t}|h_i)} = \frac{(\mu+\rho_1)\sigma_v^2 + \mu(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2}{(1+\rho_1\mu)\sigma_v^2 + (1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2} \tag{19}$$

For talent, the correlation $\mathbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i)$ is given in Appendix 1.

3.4 Summary

In Proposition 1 we summarize how a more progressive public policy (lower μ) affects the moments of income and talent.

Proposition 1. Effects of Progressivity on Income and Talent: *In any stationary state, with a time invariant public policy* $0 < \mu_{t+1} = \mu_t = \mu \leq 1$ *we have:*

[6] In (16) the variance is not indexed by i and as a result $\mathbf{E}_{h_i}(\mathbf{Var}(y_{i,t+1}|h_i)) = \mathbf{Var}(y_{i,t+1}|h_i)$. The variance of income is common to all families i because h_i enters multiplicatively into the production of talent (6). The same comment applies for the intergenerational correlation of incomes, in Section 3.3.3. In a more general version of our model, we could allow for heterogeneity in the returns to investment (e.g. with a production function of the form: $\Theta_{i,t} = (h_i V_{i,t+1}) I_{i,t}^{\xi_i}$). Under this specification the slope of the regression (δ_1) in equation (7) depends on i.

1. *A more progressive system (lower μ) decreases / increases long-run income and talent for sufficiently high / low h_i families. A more progressive system favors families with temporarily low output, $y_{i,t} < \mathbf{E}(y_{i,t+1}|h_i)$, and it favors families with temporarily low family endowment, $v_{i,t} < \rho_0$.*

2. *The dynastic variance of income, $\mathbf{Var}(y_{i,t+1}|h_i)$, and the dynastic variance of talent, $\mathbf{Var}(\theta_{i,t+1}|h_i)$, are increasing in μ. $\mathbf{Var}(y_{i,t+1}|h_i)/\mathbf{Var}(\theta_{i,t+1}|h_i)$, i.e. the intra-family ratio of intergenerational inequalities, is bounded above by 1, and is increasing in μ.*

3. *The cross sectional inequality of income $\mathbf{Var}(y_{i,t+1})$ and the cross sectional inequality of talent $\mathbf{Var}(\theta_{i,t+1})$ increase in μ. Their ratio is bounded above by 1 and also increases in μ.*

4. *The dynastic intergenerational correlation of income $\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)$ is increasing in μ. The ratio $\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)/\mathbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i)$ is smaller than 1, and increases in μ.*

Proposition 1 shows how a more progressive public policy decreases the dynastic and cross sectional inequalities of income and talent, and also decreases the within-dynasty intergenerational correlation of income. These two predictions are consistent with the general equilibrium effects of educational subsidies as derived recently in Hassler, Rodriguez Mora and Zeira (2007). They also tend to imply a positive comovement of the cross sectional and the intergenerational inequality, as discussed in Solon (2004). Finally, our model predicts that in a society with no public policy ($\mu = 1$), the ratio of variances and intergenerational correlations of income over talent take their maximum value (unity). As public policy becomes more progressive these ratios decrease. Intuitively, when the progressivity of public education increases, a given amount of variation in the production of talent across time or across families matters less for final earnings in the market.[7] In Section 5 we offer some evidence in line with this prediction.

4 The Political Economy of Social Mobility

First, we define the politico-economic equilibrium. Then, we derive the equilibrium choice of the public policy μ in terms of deeper political, economic, cultural and genetic parameters.

[7] This result reflects the difference between the coefficients δ_2 and λ_2 (or δ_i and λ_i) in the two intergenerational transmission equations. See Appendix for the details.

Finally, we show the relationship between the equilibrium level of μ and the slope of the Galton-Becker-Solon regression, β.

4.1 Politico-Economic Equilibrium

In period t, father i observes and takes as given the realization of last period's output, $y_{i,t}$, and endowment, $v_{i,t}$. However, fathers do not know the realization of children's endowment $v_{i,t+1}$ and market luck $u_{i,t+1}$ before they vote for μ_{t+1} and they need to form rational expectations. Father i's preferences over public policies μ_{t+1} are ordered according to the conditional expectation of (4):

$$W(\mu_{t+1}; h_i, y_{i,t}, v_{i,t}, s) = \ln C_{i,t} + \frac{1}{\gamma}\mathbf{E}_t(y_{i,t+1}|h_i) \tag{20}$$

where s is the vector of structural parameters, and the conditional expectation, $\mathbf{E}_t(y_{i,t+1}|h_i)$, is given by (15). $C_{i,t}$ is the optimal level of consumption:

$$C_{i,t} = \frac{\gamma}{\mu_{t+1} + \gamma} Y_{i,t} \tag{21}$$

which is a function of the public policy. Note that we reinstate the time subscript in μ.

An important simplification for deriving the equilibrium in our model is that sons are born after fathers have chosen the public policy μ_{t+1}. As a result, sons do not affect the choice of μ. Under this assumption, preferences of fathers over current policies are independent of future policies, and there is no need to explicitly consider dynasties' expectations about future policy outcomes.[8] This assumption is intuitive in the context of intergenerational mobility. As we discuss in Section 5 in a cross section of OECD countries, it is public spending on education—rather than other forms of government activity—that strongly correlates with social mobility. Since public education is regarded as highly redistributive at the primary level, i.e. before sons' political rights are extended, our assumption captures this realistic feature of the intergenerational transmission.

The policy that maximizes (20) is called the "most preferred policy for dynasty i":

$$\mu_{i,t+1} = \mu(h_i, y_{i,t}, v_{i,t}; s) = \arg\max_{\mu} W(\mu; h_i, y_{i,t}, v_{i,t}, s) \tag{22}$$

[8] That is, the indirect utility W in (20) depends only on the current choice variable, μ_{t+1}, and not on future public policies, $\mu_{t+2}, \ldots$. As a result, we do not have to consider the policy fixed point problem that arises when current policies depend on expectations of future policies but also affect future policies through the optimal consumption and investment choices and the resulting intergenerational transmission of income and talent. Our setup resembles the equilibrium in the models of Persson and Tabellini (1994), Benabou (1996), and Fernandez and Rogerson (1998), with "one period-ahead commitment to policy". Krusell, Quadrini and Rios-Rull (1997) show how to formulate and numerically solve for time-consistent politico-economic equilibria in a general class of models. Hassler, Rodriguez Mora, Storesletten and Zilibotti (2003) solve closed-form the Markov Perfect Equilibrium in a non trivial dynamic voting game under the assumption of risk neutrality.

The most preferred policy for every father reflects various trade-offs. First, it reflects the four channels that affect the long-run value of income in Section 3.3.1. In addition, transitory deviations from long-run income and transitory deviations from long-run family endowment also affect the most preferred policy, as shown in equation (15). Finally, public policy allocates resources intertemporally and creates a trade-off across generations. The consumption-investment ratio for every father is γ/μ_{t+1}. A less progressive system (higher μ_{t+1}) distorts less the incentive of parents to privately invest in their children talent and therefore when μ_{t+1} decreases parents transfer more resources to the next generation.

To solve the model we define a relevant family-specific summary of the system which we call "income potential", $Q_{i,t}$. Income potential therefore summarizes the history of all relevant market and family shocks. Our functional form assumptions—log preferences and multiplicative production functions—imply that income potential for family i at time t is the log-sum of three terms: life-long ability level $\ln h_i$, current log income, $y_{i,t}$, plus a term proportional to log family endowment, $v_{i,t}$.

$$Q_{i,t} = \ln h_i + y_{i,t} + \rho_1 v_{i,t} \tag{23}$$

Proposition 2. Preferences over Public Policy:

1. *Induced preferences over* $\mu_{i,t+1}$ *as described by* $W(.)$ *in (20) are single-peaked if (but not only if)* $\alpha > 1$ *for any* $Q_{i,t}$.

2. *The most preferred policy* $\mu_{i,t+1}$ *is strictly increasing in* $Q_{i,t}$.

The first part of the Proposition establishes a sufficient condition for the indirect utility W to be single-peaked. The second part shows that families with higher income potential prefer less progressive public policies. Families with high income potential may be families from advantaged groups (high h_i) or families that face favorable economic ($y_{i,t} > \mathbf{E}(y_{i,t+1}|h_i)$) or cultural ($v_{i,t} > \rho_0$) shocks. Therefore, in our model families from disadvantaged social groups (low h_i) may still prefer less progressive public policies, if their last generations experienced good luck in the market or in the production of talent.

Because transitory shocks affect preferences for public policies, in general the equilibrium policy will not be time invariant, as assumed for simplicity in Section 3. The easiest but most restrictive way to proceed is to assume a pre-commitment institution in which the initial generation of fathers observe $\{y_{i,0}, v_{i,0}, h_i\}$ and choose once and for all a time invariant system μ, which by assumption remains active in all future periods. A second possibility is to consider the stochastic steady state of the model, in which the distribution of income potentials in the population is stationary. In this case, the optimal μ remains constant in time, but the identity of the decisive family is allowed to vary, since in the steady state families are hit by different market and family shocks. Under both these cases, the analysis for the long-run moments in Section 3 applies, and the time invariant coefficients for the stochastic processes are given by the optimal stationary μ. Finally, we can apply our comparative statics to the most general case, when the dynastic variance and the public policy depend on calendar time along the transitional dynamics in a period-by-period decision making process. Under this setting, the equilibrium public policy (yet to be defined) will in general depend on the current state $Q_{i,t}$ of the decisive father.[9]

Let the distribution of income potential in the cross section of dynasties at time t be $\Phi_t(Q) = \int_{Q_{l,t}}^{Q} \phi_t(z)dz$. We define the political institution in terms of the equilibrium outcome that it implies.

Definition 1. Institution P: *An institution P results in the public policy μ^e_{t+1} mostly preferred by the dynasty in the $100p$th percentile of the income potential distribution Φ_t, i.e. the family with an income potential such that $p = \Phi_t(Q_{i,t})$. We denote the decisive dynasty as $Q_{p,t}$.*

Our definition encompasses some commonly used institutions, both in the optimal policy and in the political economy literature. Let the average income potential be $\bar{Q}_t = \int_{\mathbf{Q_t}} zd\Phi_t(z)$. Then if $p = \Phi_t(\bar{Q}_t)$, one obtains the utilitarian social rule that maximizes the welfare of the average father or the welfare "behind the veil of ignorance" for Q_i:

$$\max_{\mu} \int_{\mathbf{Q_t}} W(\mu, z; s)d\Phi_t(z) \tag{24}$$

In reality, however, public policies are determined by the aggregation of known, conflicting political interests. The leading choice in the political economy literature is the one person, one

[9] As a result, income and talent become regime switching stochastic processes, i.e. with time varying coefficients. One interesting and realistic case occurs if there is an adjustment cost associated with an educational reform that aims to switch μ. In this case the process for output would be a threshold ARMA(2,1) process, where the thresholds are defined by the distribution of $Q_{i,t}$ in the cross section of families. For instance, suppose that the fixed costs of expanding the public schooling infrastructure are too prohibitive and therefore μ can take only two values: $0 < \mu_l < \mu_h < 1$. Assuming that in period $t-1$, $\mu_t = \mu_h$ was the optimal grandfather's choice, a majority of fathers support a switch of regime to $\mu_{t+1} = \mu_l$, if $\int_{Q_{l,t}}^{K} \phi_t(z)dz > 1/2$ where $K = \ln\frac{\gamma+\mu_l}{\gamma+\mu_h}\left(\frac{\gamma}{\mu_l-\mu_h} - 1\right) - \ln\frac{\mu_h}{\mu_l} + \frac{\alpha}{\mu_l-\mu_h}\ln\frac{\mu_h}{\mu_l} - \rho_0(1-\rho_1)$ is a constant, ϕ_t denotes the probability distribution of income potential in the cross section of dynasties as of the beginning of period t and $Q_{l,t}$ is the lowest realized income potential. We index the distribution by t to show the possible dependency on μ_t and hence on calendar time. Under this setting, the expectations, variances and intergenerational correlations derived in Section 3 hold within each educational regime.

vote democratic institution. If $\alpha > 1$, then by Proposition 2 induced preferences over policies are single-peaked. As a result, the father with the median most preferred policy is the decisive voter. By the second part of the same Proposition, this is the father with the median income potential, $Q_{50,t}$. Note that this formulation allows both the identity, and the income or the family endowment of the decisive father to vary over time. Since the median father's vote is decisive, it follows that $p = 1/2$ is the unique equilibrium outcome of the pure majority rule game (i.e. the Condorcet winner).

More in general, we can allow for $p > 1/2$, capturing campaign contributions or more active political participation of the rich fathers. Alternatively, a higher p may parameterize the ideologically diverse preferences for parties of the poor fathers, as in the probabilistic voting model. If $p < 1/2$, then social preferences are averse to inequality and can be thought to internalize the *ex-ante* variance given in (18). From a political economy point of view, a lower p may capture the bargaining power of socialist parties or labor organizations in unionized economies. In the limit, $p = 0$ leads to the "Rawlsian institution" that maximizes the welfare of the least well-off dynasty. Henceforth, we parameterize political preferences with p. In Section 5 we show how to measure this key parameter in the data.

4.2 Politico-Economic Determinants of Social Mobility

Given this definition, the properties of the equilibrium level of the public policy, μ^e_{t+1}, are given in the following Proposition.

Proposition 3. Equilibrium Public Policy: *The equilibrium policy μ^e_{t+1} is increasing both in α and in p. It increases in h_p, $y_{p,t}$, $v_{p,t}$ and in ρ_0, it decreases in γ, and it does not depend on σ^2_v and σ^2_u. It increases in ρ_1 if and only if $v_{i,p} - \rho_0 > 0$.*

This Proposition shows how public education becomes less progressive (higher μ_{t+1}) when output costs α increase, but more progressive as the position of the decisive dynasty in the income potential distribution p decreases. Our result shows that, as long as optimally chosen public policies have the potential to affect intergenerational mobility (which in our model is shown in Section 4.3), there is no reason to expect that a collective action of fathers transmits a perfectly mobile society to their sons ($\mu^e = \beta = 0$). It is important to note that for the refusal of this proposition, one would need to show *both* that the costs of progressive public policies are negligible *and* that institutions favor the low ability families. This is an interesting point, because empirically it may be difficult to find evidence for the magnitude of α or in reality some public reforms may entail small efficiency costs (Lindert, 2004). On the other hand, a recent strand of research in political economy points out that various politico-economic outcomes can be simply explained by the fact that rich families have a larger "say" in the political equilibrium, i.e. that the political system is wealth-biased (Benabou, 1996; Campante, 2007; Alesina and Giuliano, 2009; Barenboim and Karabarbounis, 2009).

What is the novelty of our results? Most of the existing literature following the initial Becker and Tomes (1979) contribution has attributed to the reduced-form coefficient in (1) a specific meaning for social mobility, namely that equality of opportunity is desirable.[10] If equality of opportunity is however costly for private incentives, more of it is not necessarily desirable.[11] Relative to these views, our model emphasizes—in addition to standard incentive costs—political economy constraints that may further limit or enhance the extent of social mobility. For instance, in our model perfect social mobility may be optimal under a utilitarian institution (if α is very small), but *not politically sustainable* if rich families and business interests restrict the development of the welfare state and the provision of public education (i.e. if p is sufficiently high). To put it differently, two societies with similar dynastic fundamentals may display very different degrees of intergenerational mobility depending on which is the decisive dynasty selected by the existing political institutions.

The politico-economic trade-off behind our model can be conceptualized by a decline in the position of the decisive voter p. Societies in which families with lower income potential have a larger "say" for the equilibrium outcome, choose more progressive systems, expect higher mobility and lower inequality. However, progressivity results in a lower long-run level of income for sufficiently high ability families, and may even lower average income.[12] In our model, if the distribution of income potential ϕ_t is right skewed ($Q_{50} < \bar{Q}$)—perhaps because the ability distribution ϕ_h is skewed—then a majority voting of fathers chooses a more progressive public

[10]Becker and Tomes (1979; abstract and page 1182) argue that "Intergenerational mobility measures the effect of a family on the *well-being* of its children." (emphasis added). Another influential contribution is that of Mulligan (1997, page 25), who in defining social mobility notes that "The degree of intergenerational mobility is [...] an index of the degree of 'equality of opportunity'. Equality of opportunity is often seen as *desirable* because, with little correlation between the incomes of parents and children, children from rich families do not enjoy much of a 'head start' on children from poor families.". The same presumption may be implied by the introductory paragraph in Solon (1999).

[11]Piketty (2000) and Corak (2006) make this point. In an influential paper, Atkeson and Lucas (1992) have shown the optimality of zero mobility. Recently, Phelan (2006) and Farhi and Werning (2007) challenge this result based on the social discount rate exceeding the private one.

[12]We have not explicitly considered the growth enhancing effects of public education. However, if average ability $\bar{h}$ is sufficiently low, then in the steady state the stationary average income in the cross section of the dynasties, $\int_{\mathbf{H}} \mathbf{E}(y_{i,t+1}|h) d\Phi_h(h)$, is decreasing in μ, and the progressivity increases long-run income, which implicitly may be capturing this realistic feature of public education.

policy relative to the utilitarian optimum. Holding average income potential $\bar{Q}$ constant, an increase in the (right) skewness of the distribution of income potentials, leads the majority of fathers to demand more progressive policies and higher social mobility.

Interestingly, the effects of a higher *ex-ante* inequality in abilities, $\mathbf{Var}(\ln h)$, due for instance to market discrimination against ethnically or racially diverse groups, depend on the political process p. If p is low, then higher $\mathbf{Var}(\ln h)$ could be associated with more skewness and hence a poorer decisive voter which results in more progressive policies. On the other hand, if *de facto* political power is ultimately related to income potential and hence p is relatively high, a higher *ex-ante* variability could be associated with more powerful elites, less progressivity and lower social mobility. In Section 5 we offer some suggestive evidence in favor of the second effect. In the Appendix we discuss in more detail the intuition behind the other comparative statics of our model.

4.3 Structural Politico-Economic Interpretation of the Galton-Becker-Solon (GBS) Regression

Our theoretical framework offers a structural interpretation for the log-linear intergenerational earnings model which is estimated in the empirical literature cited in Section 2. The literature typically focuses on the Galton-Becker-Solon (GBS) regression:

$$y_{i,t+1} = a + \beta y_{i,t} + \varepsilon_i \tag{25}$$

where y_{t+1} and y_t denote son's and father's life-long log earnings in the population. Previous models have recognized that β is a function of genetic and cultural inheritance, altruism, technological parameters and the structure of the asset market. However, we show that this coefficient also depends on political economy variables which determine the institutions that a generation puts in place to insure its offspring from adverse shocks.

Proposition 4. Population Slope of the GBS Regression: *The slope in the population regression of son's on father's income, β, also known as the intergenerational elasticity of income is given as follows.*[13]

1. *If the economy is in a stationary state with $\mu^e_{t+1} = \mu^e_t = \mu^e$, then the intergenerational elasticity equals the intergenerational correlation of incomes and is given by:*

$$\beta = \mathbf{Corr}(y_{i,t+1}, y_{i,t}) = \mu^e \left(1 + \frac{\frac{\rho_1 \mu^e}{(1-\rho_1^2)(1-\rho_1\mu^e)}\sigma_v^2 + \frac{\mu^e}{1-\mu^e}\mathbf{Var}(\ln h_i)}{\mathbf{Var}(y_{i,t};\mu^e)}\right) \tag{26}$$

[13]Note that in both cases β is expressed only as a function of the deeper parameters of the model.

where the variance in the denominator refers to the cross sectional variance in (17) and μ^e is the equilibrium public policy defined in Proposition 3. The intergenerational elasticity β increases in μ^e and in p.

2. *If the economy is for a long time in the steady state $\mu_t^e = \mu_{t-1}^e = ...$, but in $t+1$ an unexpected structural break in the political institution p happens, then the intergenerational elasticity is given by:*

$$\beta_{t+1} = \mu_{t+1}^e \left(1 + \frac{\frac{\rho_1 \mu_t^e \sigma_v^2}{(1-\rho_1^2)(1-\rho_1 \mu_t^e)} + \frac{\mu_t^e}{1-\mu_t^e} \mathbf{Var}(\ln h_i)}{\mathbf{Var}(y_{i,t}; \mu_t^e)} \right) \tag{27}$$

In this case $\beta_{t+1}/\beta_t = \mu_{t+1}^e/\mu_t^e$ and the ratio is increasing in p_{t+1}/p_t.

The first part of the proposition refers to the special case in which the economy is in an steady state with constant intergenerational mobility (the coefficient β) and cross-sectional variances. The second part considers instead the case in which a political shock at time t determines a change of the decisive dynasty such that intergenerational mobility changes with respect to previous periods ($\beta_{t+1} \neq \beta_{t-i}$ for $i \geq 0$) and cross-sectional income variances may differ across generations. In principle, analogous formulas can be obtained for other shocks affecting the intergenerational elasticity of incomes and cross-sectional variances, but given the focus of this paper here we study the case of a political shock.

Under the assumption that the advanced economies for which an estimate of β is available are essentially characterized by a fairly similar set of economic and biological fundamentals, differences in the estimated β for these countries should correlate with differences in the dynasty that has decisive power in the political process. To put it differently, if economic and biological fundamentals are more similar than political equilibria across these advanced economies, we should observe more mobility in countries in which the position of the decisive dynasty is lower in the hierarchy of dynastic income potentials. The empirical exercise in the next section should be interpreted as a suggestive assessment of the extent to which political economy variables that proxy for the decisive dynasty capture the cross-country variation in β.

However, equations (26) and (27) emphasize also other more traditional determinants that might explain the cross country variability in β. For example, in steady state and for given decisive dynasty, social mobility increases (β decreases) with market luck variability (higher σ_u^2), and decreases with *ex-ante* heterogeneity (higher $\mathbf{Var}(\ln h_i)$). It decreases with output costs (higher α), with the ability of the decisive family (higher Q_p), with the long-run family endowment (higher ρ_0) and with the degree of altruism (higher γ). Greater market variability increases cross sectional inequality and makes the position of children highly uncertain, thereby increasing social mobility. For the same reason, the comparative static with respect to σ_v^2 and ρ_1 is theoretically ambiguous. α, p, h_p, ρ_0, and γ affect social mobility indirectly, through the equilibrium level of μ (see Appendix for these comparative statics). Finally, note that our model predicts that *ex-ante* heterogeneity $\mathbf{Var}(\ln h_i)$ affects positively β only conditional on μ. Higher *ex-ante* heterogeneity may operate also indirectly through public policy, and it may increase (if it is associated with smaller p) or decrease (under higher p) social mobility.

5 Empirical Evidence on the Politico-Economic Determinants of Mobility

In this Section we turn to the empirical evidence on the predictions of the model. Specifically, we present evidence on the relationship between political variables that our model indicates as important determinants of social mobility and observed measures of mobility across countries or within country over time. We stress that this evidence needs to be interpreted only as suggestive and descriptive. The number of countries for which we have data is limited, and the available data are not sufficiently informative to identify causal relationships. Nevertheless, the evidence is generally consistent with the predictions of our model and in particular, it supports a positive cross country and within country correlation between proxies for p and estimates of β.

We consider first an interesting case study which represents a salient example of a political shock as described in equation (27). Over the past few decades, the UK has experienced a tremendous decline in social mobility. In particular, Blanden, Gregg and Macmillan (2007) document a 50% decline in social mobility between the 1958 and the 1970 cohorts. Such decline has generated widespread concern among the public and has prompted the government in 2009 to issue a White Book that addresses the causes and implications of the decline in mobility. Blanden, Gregg and Macmillan (2007) argue that the main cause of the decline is represented by changes in educational attainment of different income groups.

Their evidence is consistent with our model. However, our model goes further and indicates that educational policies are likely to be an endogenous outcome. According to our framework, the ultimate determinant of the decline in social mobility should be a change in societal preferences for redistribution. Indeed, this prediction is consistent with the sharp change in political preferences which led Margaret Thatcher to become Prime Minister in 1979. The Thatcher revolution was caused by a clear move toward the right by the UK electorate, as indicated by the fact that public expenditure for education fell, the power of the unions declined, regressive VAT taxes increased, and more progressive corporate and income taxes declined.[14]

Turning to cross country evidence, credible estimates of β are available only for a limited number of countries. We use estimates from Corak's (2006) meta-analysis conducted for 9 OECD countries and complement these with 3 more observations. In the Appendix we discuss more in detail the construction of our dataset and the sources.[15]

In Figure 2, the vertical axis shows estimates of β for a cross section of advanced democratic OECD countries. Consistent with what has long been documented in the existing literature on mobility, the UK, US and France are the least mobile, while Northern European countries appear the most mobile. Canada is the most mobile Anglo-Saxon country, and Sweden is the least mobile among the Nordic countries. The existing literature has mostly focused on the left panel of the Figure (for example: Corak, 2006), which shows a positive bivariate association between β and the private return to schooling. The right panel, which is more novel, depicts a negative association between β and public expenditure on education. The Figure shows that the correlation between social mobility and public expenditures for education is at least as strong as the correlation between the private internal return to education and mobility.[16] When we divide public expenditure in education per student as a percentage of per capita GDP at the primary, secondary and tertiary level, we find that all are negatively correlated with β. Notably and consistent with our model, the correlation is stronger at the primary level, where public expenditures are arguably more redistributive.[17]

To obtain a direct measure of political preferences (the parameter p in the model), we

[14] VAT taxes rose around 15%, and each of the corporate tax rate and the top marginal income tax decreased by 17%. Public expenditure for education as a percentage of GDP decreased by 25% between 1975 and 1985 and by 30% by the end of the 1980s.

[15] In the following Figures we use Corak's most preferred estimate, but we have verified the robustness of our results using the median estimate found in the literature. The nine countries are Denmark, Norway, Finland, Canada, Sweden, Germany, France, US and the UK. We also add Japan, Spain and Australia. Some recent papers have estimated the intergenerational income elasticity in Italy, but: (i) the estimates are based on heroic assumptions needed to use intergenerational income data of low quality; (ii) the estimates are especially high and (iii) even using a more conservative value, Italy is most of the times a major outlier of which we cannot be really confident. The only variable that seems to explain satisfactory Italy's low degree of mobility is the strength of family ties (high ρ_1).

[16] Conditioning on both determinants, the latter turns out to be much more strongly associated with mobility than the former (correlation of -0.43 versus 0.15).

[17] In contrast, the correlation of β with total government spending is -0.05, and the correlation of β with spending on social expenditures is -0.11. The weakness of these correlations illustrates that it is educational expenditure, rather than other forms of government spending (e.g. unemployment insurance, assistance to poor families, welfare benefits, etc.), that may matter for social mobility.

use data from the World Value Surveys (WVS).[18] We focus on the differential in political participation between low income voters and middle and upper income voters. The income classification follows the WVS and is standardized by country. As Table 1 shows, on average, around 33% of the population is classified as "poor" (low-income). Variation across countries is not large.[19] Political participation can be measured with a variety of variables. In Figure 3 we measure political participation with membership in political parties. The vertical axis in the Figure measures inequality in party affiliation, defined as the fraction of middle and upper income voters who are members of political parties divided by the fraction of low income voters who are members of political parties. A lower value for this index denotes a relatively more politically active class of low income families and hence a lower p. Note that we are not interested in the political participation of the poor per se, but in their participation *relative* to that one of other income groups in the same country. Our measure of relative participation therefore holds constant other country-specific factors that may affect political participation.

The correlations in Figure 3 are consistent with the model. The bivariate correlations of the political inequality index with public spending and intergenerational elasticity are, respectively, -0.49 and 0.79.[20] When we use an alternative measure of the gap in political participation that compares participation by high income voters to participation by with low income voters (thus excluding middle income voters), the correlation is even stronger.

The existing literature has argued that one of the most important empirical determinants of social mobility is the rate of return to human capital (see for example Solon, 1999 and 2004; Corak, 2006). When we regress β on estimates of the return to schooling, we find that the return to schooling explains only 8% of the cross country variation. Notably, and consistent with our model, our measure of inequality between rich and the poor families in political affiliation explains 42% of the variation in social mobility.[21] We have repeated this exercise with four other measures of political participation: participation in labor unions, interest in politics, signing petitions and participating in lawful demonstrations. We find that the patterns are similar to those presented, with the bivariate correlations ranging from 0.43 to 0.63. (Results available upon request.)

[18]In a previous version of the paper we used voter turnout in elections and union density as additional proxies for p. For all cases we find correlations between p, μ and β that are consistent with our model.

[19]Sweden and Germany are the two outliers.

[20]This finding is robust to the exclusion of outliers.

[21]One of the few studies that attribute cross country differences in mobility to public policies is Corak and Heitz (1999). The authors conjecture that Canada's progressivity can explain its higher mobility relative to the US.

In Figure 4 we investigate the relationship between the degree of heterogeneity in a society, public education and social mobility. Our model predicts that higher *ex-ante* heterogeneity (higher $\mathbf{Var}(\ln h_i)$) should be associated with more public spending and therefore higher social mobility if p is low. If p is high, more heterogeneity should be associated with higher talent (h_p) for the decisive family, and less progressivity. Our empirical proxy for heterogeneity is an index of ethnolinguistic fragmentation measured in 1961.[22] The upper left panel shows that more diverse countries are associated with less public spending on education. Our model explains this positive correlation only if p is relatively high, which as discussed above is consistent with recent theoretical and empirical literature. The bottom panel shows that the predicted link between heterogeneity and mobility is also supported by the data. The bivariate correlation is 0.26. Excluding the very heterogeneous and mobile Canada, the correlation increases to 0.67.

Another prediction of the model has to do with the strength of cultural transmission ρ_1. As a proxy, we use an index of weak family ties.[23] Weaker family ties proxy for a lower ρ_1 in our model. In Figure 4, weaker family ties are associated with more public provision of education and more mobility. This lends support to the view that strong family ties and strong social policies are substitutes.

We conclude with a final piece of evidence. Becker and Tomes (1979) original contribution aimed at explaining within a unified economic model the degree of cross sectional inequality, and its relation with intergenerational inequality. We proxy for cross sectional inequality in earnings, $\mathbf{Var}(y_{i,t})$, with the Gini coefficient for *gross* earnings. The variance in talent or skills, $\mathbf{Var}(\theta_{i,t})$, is proxied by the Gini coefficient for *factor* income.[24] In our sample the bivariate association between cross sectional gross earnings inequality and intergenerational inequality is around 0.72. Within the context of our model, market variability, σ_u^2, explains the lack of perfect correlation. Higher variability increases cross sectional inequality to a degree that ultimately raises social mobility.[25]

Proposition 1 implies that the ratio of gross earnings over factor inequality should decline when the progressivity of the educational system increases (μ decreases). Figure 5 shows a strong association between the ratio of the Gini coefficients and public expenditure in educa-

[22] The index is defined as one minus the probability that two random persons in some country belong to the same ethnic, linguistic or racial group.

[23] The index is due to Alesina and Giuliano (2007). We thank the authors for providing us with their data.

[24] These statistics come from Milanovic (2000).

[25] Björklund and Jäntti (1997) hypothesize that common causes may explain US's higher intergenerational and cross sectional inequality relative to Sweden's. Recently, Hassler, Rodriguez Mora and Zeira (2007) argue that inequality and mobility may be positively correlated if labor market institutions differ significantly across countries or negatively correlated if educational subsidies drive the cross country variation.

tion. It also shows the direct relationship between the deeper determinant p and the ratio of inequalities $\mathbf{Var}(y_{i,t})/\mathbf{Var}(\theta_{i,t})$ that can rationalize this association. In particular, our model predicts that in societies where the poor participate more in political parties, redistributive public education takes place and therefore the ratio of income over talent inequality decreases. The right panel of the figure is consistent with this prediction.

6 Conclusion

Intergenerational mobility emerges from "nature", "nurture" *and* endogenous public policies. While the previous literature has derived social mobility as a function of the optimizing behavior of utility-maximizing families, in this paper we generalize the structural log-linear social mobility model and endogenize the political process that aggregates conflicting preferences for intergenerational mobility.

Our model provides a structural interpretation of the widely-studied Galton-Becker-Solon reduced-form coefficient β. This is important because it allows a better understanding of the deeper economic determinants of intergenerational mobility and the role of public policy. We show that public policies generate a trade-off between insurance and incentives. Our model adds to this knowledge by pointing out that even if insurance is relatively costless to provide, a less than perfectly mobile society is possible because of political economy constraints in a world of heterogeneous interests. In other words, two societies may have the same set of dynastic fundamentals such as parental altruism, level of GDP, asset markets, ethnic fragmentation and cultural traits, but different political institutions, in which case social mobility outcomes will differ.

We conclude with some empirical evidence that lends support to our claim that politico-economic variables are likely to be important determinants of cross country differences in social mobility.

References

Aaronson, Daniel and Bhashkar Mazumder. 2008. "Intergenerational Economic Mobility in the US: 1940 to 2000." *Journal of Human Resources,* 43(1): 139–172.

Alesina, Alberto, and Paola Giuliano. 2007. "The Power of the Family." NBER Working Paper, No. 13051.

Alesina, Alberto, and Paola Giuliano. 2009. "Preferences for Redistribution." NBER Working Paper, No. 14825.

Alesina, Alberto, and Edward Glaeser. 2004. *Fighting Poverty in the US and Europe: A World of Difference.* Oxford, UK: Oxford University Press.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics,* 109: 465–490.

Atkeson, Andrew, and Robert Lucas. 1992. "On Efficient Distribution with Private Information." *Review of Economic Studies,* 59: 427–453.

Barenboim, Igor, and Loukas Karabarbounis. 2009. "One Dollar, One Vote." mimeo, Harvard University.

Becker, Gary, and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy,* 87: 1153–1189.

Becker, Gary, and Nigel Tomes. 1986. "Human capital and the rise and fall of families." *Journal of Labor Economics,* 4: S1–S39.

Benabou, Roland. 1996. "Inequality and Growth." In *NBER Macroeconomics Annual.* ed. Ben Bernanke and Julio Rotemberg, 11–73. Cambridge, MIT Press.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review,* 90: 96–129.

Benabou, Roland, and Efe Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *Quarterly Journal of Economics,* 116(2): 447–487.

Bernasconi, Michele, and Paola Profeta. 2007. "Redistribution or Education? The Political Economy of the Social Race." CESifo Working Paper No. 1934.

Björklund, Anders, and Markus Jäntti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review,* 87: 1009-1018.

Blanden, Jo, Paul Gregg, and Lindsey Macmillan. 2007. "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education." *Economic Journal,* 117(1): C43–C60.

Boarini, Romina, and Hubert Strauss. 2007. "The Private Internal Rates of Return to Tertiary Education: New Estimates for 21 OECD Countries." *OECD Economics Department Working Papers,* No. 591.

Campante, Filipe. 2007. "Redistribution in a Model of Voting and Campaign Contributions." mimeo, Harvard University.

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini. 1999. "More Equal But Less Mobile? Education Financing and Intergenerational Mobility in Italy and the US." *Journal of Public Economics,* 74: 351–393.

Corak, Miles, and Andrew Heitz. 1999. "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data." *Journal of Human Resources,* 34(3): 504–533.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons From a Cross Country Comparison of Generational Earnings Mobility." *Research on Economic Inequality,* 13(1): 143—188.

d'Addio, Anna Christina. 2007. "Intergenerational Transmission of Disadvantage: Mobility or Immobility Across Generations?" *OECD Social Employment and Migration Working Papers,* No. 52.

Davies, James, Jie Zhang, and Jinli Zeng. 2005. "Intergenerational Mobility under Private vs. Public Education." *Scandinavian Journal of Economics*, 107(3): 399–417.

Farhi, Emmanuel, and Ivan Werning. 2007. "Inequality and Social Discounting." *Journal of Political Economy,* 115(3): 365–402.

Fernandez, Raquel, and Richard Rogerson. 1998. "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform." *American Economic Review,* 88(4): 813–833.

Glomm, Gerhard, and Balasubrahmanian Ravikumar. 1992. "Public vs. Private Investment in Human Capital: Endogenous Growth and Income Inequality." *Journal of Political Economy,* 100: 818–834.

Goldberger, Arthur. 1989. "Economic and Mechanical Models of Intergenerational Transmission." *American Economic Review,* 79(3): 504–513.

Hassler, John, Jose Vicente Rodriguez Mora, Kjetil Storesletten, and Fabrizio Zilibotti. 2003. "The Survival of the Welfare State." *American Economic Review,* 93(1): 87–112.

Hassler, John, Jose Vicente Rodriguez Mora, and Joseph Zeira. 2007. "Inequality and Mobility." *Journal of Economic Growth,* 12: 235–259.

Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull. 1997. "Politico-Economic Equilibrium and Economic Growth." *Journal of Economic Dynamics and Control,* 21: 243–272.

Lee, Chul-In, and Gary Solon. 2006. "Trends in Intergenerational Income Mobility." NBER Working Paper, No. 12007.

Lefranc, Arnaud, Fumiaki Ojima, and Takashi Yoshida. 2008. "The Intergenerational Transmission of Income and Education: A Comparison of Japan and France." Working Paper, EUI.

Leigh, Andrew. 2007. "Intergenerational Mobility in Australia." *B.E. Journal of Economic Analysis and Policy,* 7(2), Article 6.

Lindert, Peter. 2004. *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century.* Volumes 1 and 2. Cambridge University Press.

Loury, Glenn. 1981. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica,* 49(4): 843–847.

Mayer, Susan, and Leonard Lopoo. 2005. "Has the Intergenerational Transmission of Economic Status Changed?" *Journal of Human Resources,* 40(1): 169—185.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the U.S. Using Social Security Earnings Data." *Review of Economics and Statistics,* 87(2): 235–255.

Mazumder, Bhashkar. 2007. "Trends in Intergenerational Mobility." *Industrial Relations,* 46(1): 1–6.

Milanovic, Branco. 2000. "The Median Voter Hypothesis, Income Inequality and Income Redistribution: An Empirical Test with the Required Data." *European Journal of Political Economy,* 16: 367–410.

Mulligan, Casey. 1997. *Parental Priorities and Economic Inequality.* The University of Chicago Press.

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr. 2008. "School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform," mimeo, Helsinki School of Economics.

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review,* 84(3): 600–621.

Persson, Torsten, and Guido Tabellini. 2003. *The Economic Effects of Constitutions*, Cambridge, Mass: MIT Press.

Phelan, Christopher. 2006. "Opportunity and Social Mobility." *Review of Economic Studies,* 73: 487–504.

Piketty, Thomas. 1995. "Social Mobility and Redistributive Politics." *Quarterly Journal of Economics,* 110(3): 551–584.

Piketty, Thomas. 2000. "Theories of Persistent Inequality and Intergenerational Mobility." In *Handbook of Income Distribution.* ed. Antony B. Atkinson and Francois Bourguignon, Volume 1: 429-476. North-Holland.

Saint-Paul, Gilles, and Thierry Verdier. 1993. "Education, Democracy and Growth." *Journal of Development Economics,* 42, 399–407.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review,* 82: 393–408.

Solon, Gary. 1999. "Intergenerational mobility in the Labor Market." In *Handbook of Labor Economics.* ed. Orley C. Ashenfelter and David Card, Volume 3B: 1762–1796. New York, NY: Elsevier Science Press.

Solon, Gary. 2002. "Cross-Country Differences in Intergenerational Earnings Mobility." *Journal of Economic Perspectives,* 16(3): 59–66.

Solon, Gary. 2004. "A Model of Intergenerational Mobility Variation over Time and Place." In *Generational Income mobility in North America and Europe.* ed. Miles Corak, 38–47. Cambridge University Press.

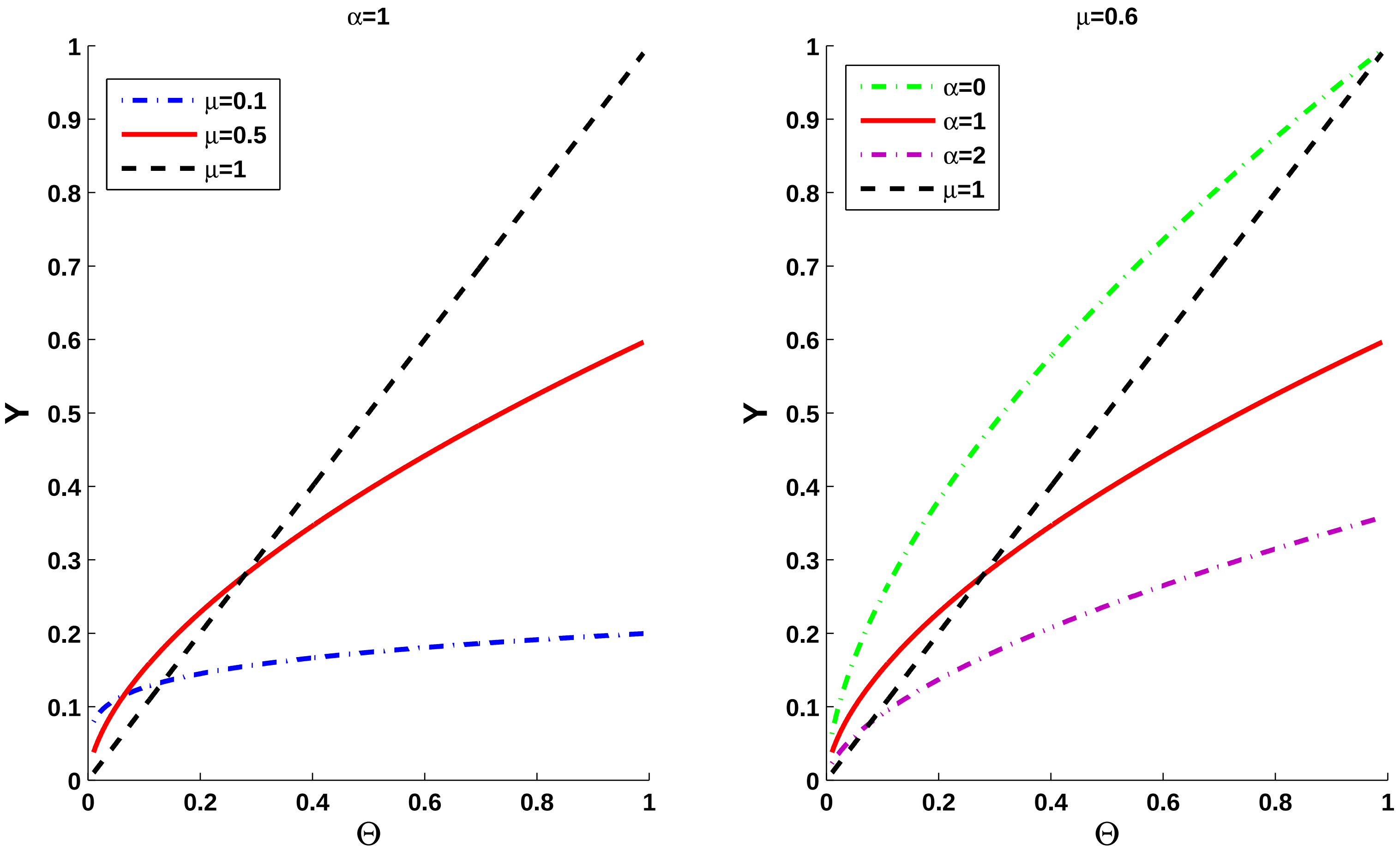
Zimmerman, David. 1992. "Regression Toward Mediocrity in Economic Stature." *American Economic Review,* 82: 409–429.

Table 1: Classification in Poor, Middle and Rich Per Country: WVS Data

	% Poor	% Middle	% Rich
Australia	29	34	37
Canada	31	36	33
Denmark	31	41	28
Finland	33	33	34
France	33	37	30
Germany	39	33	28
Japan	32	36	32
Norway	35	40	25
Spain	30	44	26
Sweden	26	44	31
UK	35	35	30
US	35	36	29
Average	33	37	30

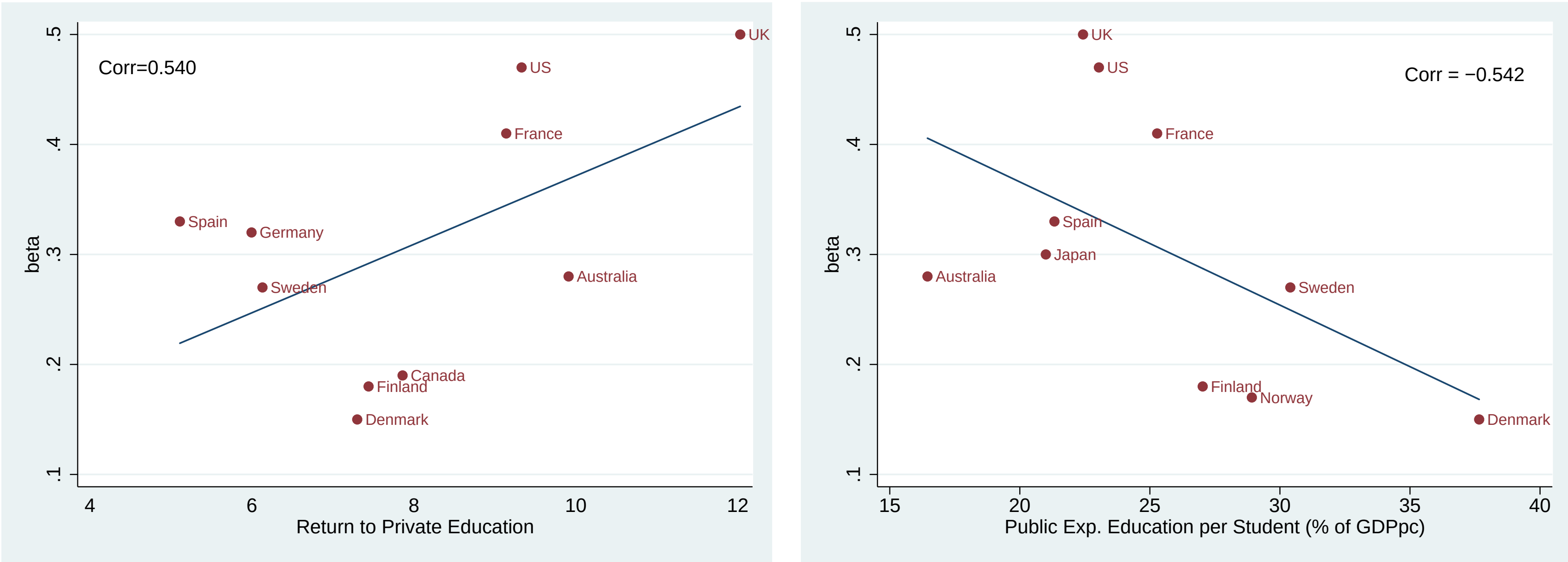
Notes: Percentages are rounded to sum to 100. The numbers refer to the full sample from the Four Wave WVS Data. Actual percentages used in the empirical results may differ slightly depending on the political variable used.

Figure 1: The Production Function $Y_{i,t} = \mu_t^{\alpha} \left(U_{i,t}\Theta_{i,t}\right)^{\mu_t}$



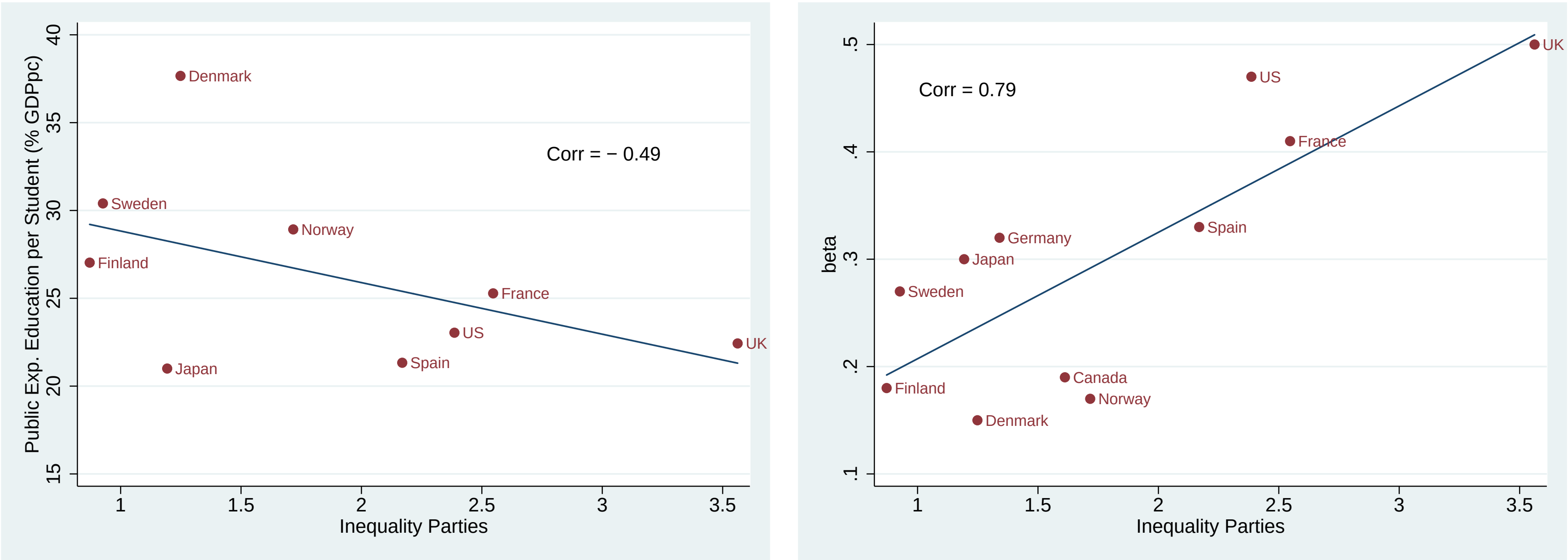
Notes: Market luck is set to $U_{i,t} = 1$.

Figure 2: Private Return to Education vs. Public Expenditure in Education



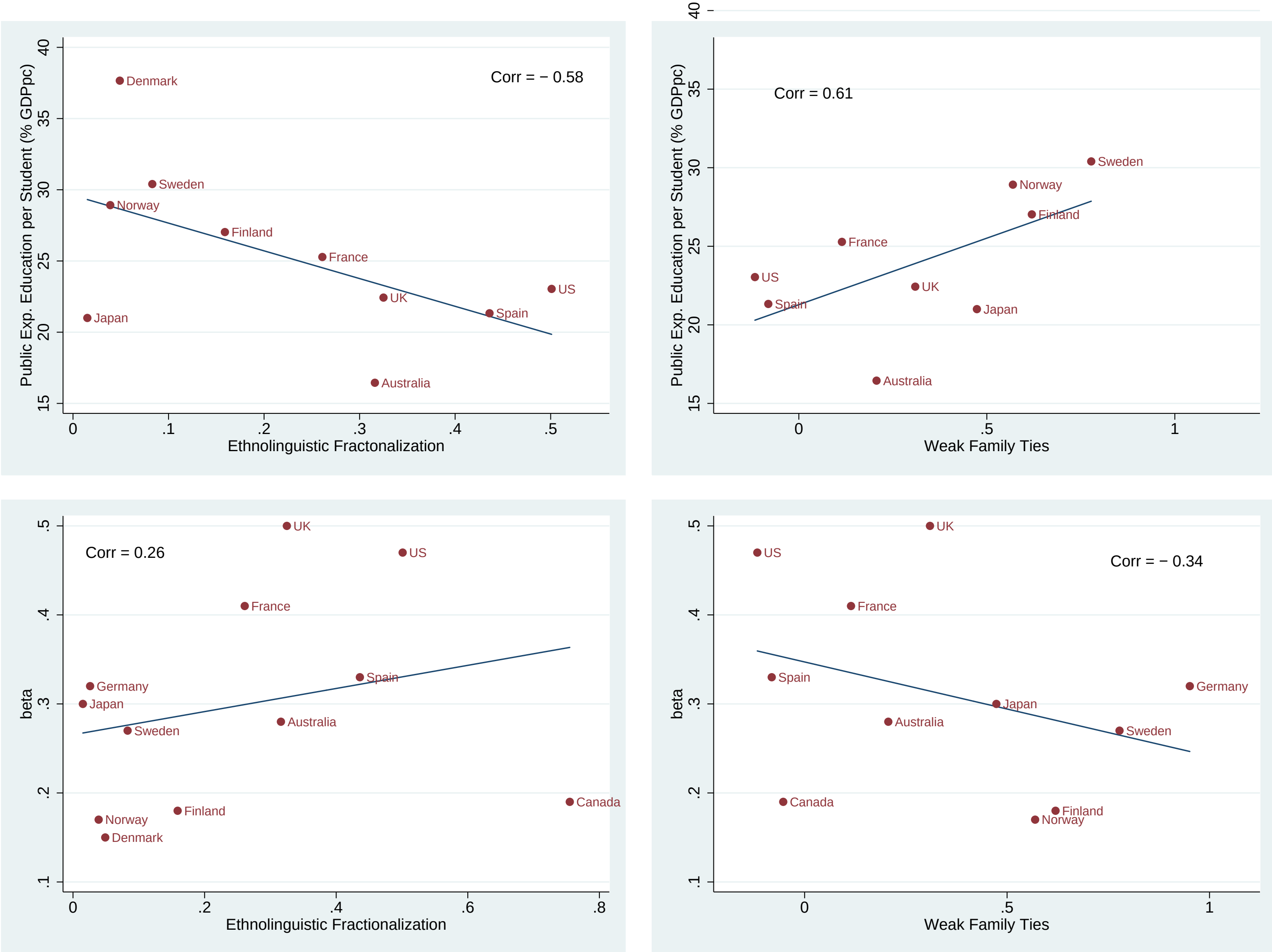
Notes: The left panel shows the relationship between the intergenerational earnings elasticity β and the private rate of return to tertiary education. The right panel shows the relationship between the intergenerational earnings elasticity β and the public expenditure in education per student as a percentage of per capita GDP. See Appendix 2 for the data sources.

Figure 3: Mobility, Public Education, and Inequality in Political Participation



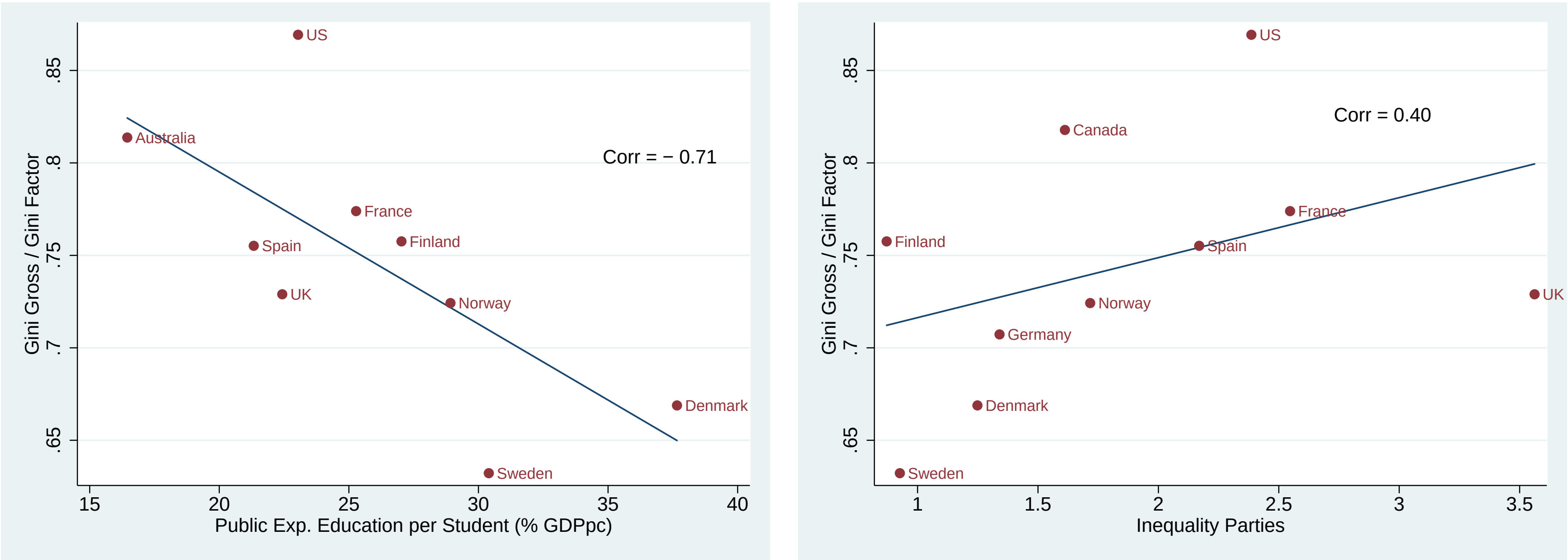
Notes: The left panel shows the relationship between the public expenditure in education per student as a percentage of per capita GDP and the variable "Inequality Parties". The right panel shows the relationship between the intergenerational earnings elasticity β and the variable "Inequality Parties". The variable "Inequality Parties" (proxy for p) is defined as the political party participation rate of the non-poor (middle and high income) citizens divided by the political party participation rate of the poor citizens. See Appendix 2 for the data sources.

Figure 4: Mobility, Public Education, Heterogeneity and Family Ties



Notes: The left panels show the relationship between the ethnolinguistic fractionalization index, the public expenditure in education per student as a percentage of per capita GDP (upper left), and the intergenerational earnings elasticity β (lower left). The right panels show the relationship between the weakness of family ties, the public expenditure in education per student as a percentage of per capita GDP (upper right), and the intergenerational earnings elasticity β (lower right). See Appendix 2 for the data sources.

Figure 5: Income and Talent Cross Sectional Inequality



Notes: The left panel shows the relationship between the ratio of Gini coefficients measured at the gross and the factor level and the public expenditure in education per student as a percentage of per capita GDP. The right panel shows the relationship between the ratio of Gini coefficients measured at the gross and the factor level and the variable "Inequality Parties". The variable "Inequality Parties" (proxy for p) is defined as the political party participation rate of the non-poor (middle and high income) citizens divided by the political party participation rate of the poor citizens. See Appendix 2 for the data sources.

Appendix 1: Derivations and Proofs

A1. Derivation of Income and Talent Transmission Equations

First, forward the production function for output, equation (2), one period and solve for $\Theta_{i,t+1}$:

$$\Theta_{i,t+1} = (\mu_{t+1})^{-\frac{\alpha}{\mu_{t+1}}} (U_{i,t+1})^{-1} (Y_{i,t+1})^{\frac{1}{\mu_{t+1}}} \tag{A.1}$$

Substitute (A.1) into the production function for talent, (6), and solve for investment:

$$I_{i,t} = (h_i V_{i,t+1})^{-1} \left[(Y_{i,t+1})^{\frac{1}{\mu_{t+1}}} (U_{i,t+1})^{-1} (\mu_{t+1})^{-\frac{\alpha}{\mu_{t+1}}} \right] \tag{A.2}$$

If we insert this equation into the budget constraint, $C_{i,t} = Y_{i,t} - I_{i,t}$, we see that the budget is concave for $\mu_{t+1} \leq 1$, strictly when $\mu_{t+1} < 1$. Since the utility function (4) is strictly concave, the solution to the problem is unique and interior and is characterized by the first order condition:

$$\frac{C_{i,t}}{\gamma Y_{i,t+1}} = \frac{1}{\mu_{t+1}(h_i V_{i,t+1}) U_{i,t+1}} (\mu_{t+1})^{-\frac{\alpha}{\mu_{t+1}}} (Y_{i,t+1})^{\frac{1}{\mu_{t+1}}-1} \tag{A.3}$$

Substituting $C_{i,t}$ back in the budget constraint, we take the solution for children's income:

$$Y_{i,t+1} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right)^{\mu_{t+1}} (h_i V_{i,t+1} U_{i,t+1})^{\mu_{t+1}} (\mu_{t+1})^{\alpha} (Y_{i,t})^{\mu_{t+1}} \tag{A.4}$$

Taking logs and letting $\mu_{t+1} = \mu$ in (A.4) yields the income transition equation (7) in the text, for the coefficients defined in (8)-(13). From (A.2) and the budget constraint we can also take the solution for investment and consumption:

$$I_{i,t} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right) Y_{i,t} \tag{A.5}$$

$$C_{i,t} = \left(\frac{\gamma}{\mu_{t+1} + \gamma} \right) Y_{i,t} \tag{A.6}$$

To derive the intergenerational transmission equation for talent, we first substitute the production function (2) into the solution (A.4). This yields a relationship between sons' income and fathers' talent:

$$Y_{i,t+1} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right)^{\mu_{t+1}} (h_i V_{i,t+1} U_{i,t+1})^{\mu_{t+1}} (\mu_{t+1})^{\alpha} \left[\mu_t^{\alpha} \Theta_{i,t}^{\mu_t} U_{i,t}^{\mu_t} \right]^{\mu_{t+1}} \tag{A.7}$$

Next, substitute (A.7) into (A.1) to obtain the solution for talent:

$$\Theta_{i,t+1} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right) (h_i V_{i,t+1}) (\mu_t)^{\alpha} U_{i,t}^{\mu_t} \Theta_{i,t}^{\mu_t} \tag{A.8}$$

Taking logs and setting $\mu_{t+1} = \mu_t = \mu$ gives the transmission equation for talent:

$$\theta_{i,t+1} = \lambda_{0,i} + \lambda_1 \theta_{i,t} + \lambda_2 v_{i,t+1} + \lambda_3 u_{i,t} \tag{A.9}$$

where:

$$\lambda_{0,i} = \lambda_0 + \lambda_i \tag{A.10}$$

$$\lambda_0 = \ln\left(\frac{\mu}{\mu+\gamma}\right) + \alpha \ln \mu \tag{A.11}$$

$$\lambda_i = \ln h_i \tag{A.12}$$

$$\lambda_1 = \mu \tag{A.13}$$

$$\lambda_2 = 1 \tag{A.14}$$

$$\lambda_3 = \mu \tag{A.15}$$

The talent transmission equation differs from the income transmission equation due to the coefficients λ_2 and λ_i (as opposed to the coefficients δ_2 and δ_i in the text). These coefficients measure the effects of cultural and genetic endowment on talent and output, respectively. For the case of talent, these effects do not depend on μ, since public policies are imposed on final output.

A2. Expected Income and Talent

First we show that given a stationary μ, income and talent are stationary processes. Subtracting $\rho_1 y_{i,t}$ from both sides of the income transmission equation (7), using the definition for $v_{i,t+1}$ in (3), and substituting in the resulting expression the fact that $\rho_1 (\delta_2 v_{i,t} - y_{i,t}) = -\rho_1 (\delta_{0,i} + \delta_1 y_{i,t-1} + \delta_3 u_{i,t})$, we can express the income process in (7) as the sum of an ARMA(2,1) process plus an independent white noise:

$$y_{i,t+1} = (1-\rho_1)(\delta_{0,i} + \delta_2 \rho_0) + (\delta_1 + \rho_1) y_{i,t} + (-\delta_1 \rho_1) y_{i,t-1} + \delta_3 u_{i,t+1} - \delta_3 \rho_1 u_{i,t} + \delta_2 \epsilon_{i,t+1} \tag{A.16}$$

The process is stationary if the roots of the characteristic equation, $1 - (\delta_1 + \rho_1)x - (-\delta_1 \rho_1)x^2 = 0$, lie outside the unit circle. The two roots are given by $\phi_1 = -\frac{1}{\rho_1}$ and

$\phi_2 = -\frac{1}{\delta_1} = -\frac{1}{\mu}$. Therefore, the log income process is stationary for every family i, if $\rho < 1$ and $\mu < 1$. A similar reasoning applies for the talent process.

The unconditional expectation of log income for family i in equation (14) in the text is easy to compute by setting $\mathbf{E}(y_{i,t+1}) = \mathbf{E}(y_{i,t}) = \mathbf{E}(y_{i,t-1})$ in (A.16) or (7). All comparative statics for this expectation follow from inspection. From the talent transmission equation (A.9), we take the unconditional expectation of log talent for family i:

$$\mathbf{E}(\theta_{i,t+1}|h_i) = \frac{\rho_0 + \ln\left(h_i \frac{\mu}{\mu+\gamma}\right) + \alpha \ln \mu}{1-\mu} \tag{A.17}$$

From the income transmission equation (7) we can compute the conditional expectation of income:

$$\mathbf{E}_t(y_{i,t+1}|h_i) = \mathbf{E}(y_{i,t+1}|h_i) + \mu\left(y_{i,t} - \mathbf{E}(y_{i,t+1}|h_i)\right) + \mu\rho_1\left(v_{i,t} - \rho_0\right) \tag{A.18}$$

where the state of the system includes $\{y_{i,t}, \theta_{i,t}, v_{i,t}, u_{i,t}\}$, and $\mathbf{E}(y_{i,t+1}|h_i)$ is the unconditional expectation given in (14). Similarly for talent we have:

$$\mathbf{E}_t(\theta_{i,t+1}|h_i) = \mathbf{E}(\theta_{i,t+1}|h_i) + \mu\left(\theta_{i,t} - \mathbf{E}(\theta_{i,t+1}|h_i)\right) + \rho_1\left(v_{i,t} - \rho_0\right) \tag{A.19}$$

A3. Variance of Income and Talent

To derive the unconditional, stationary variance $\mathbf{Var}(y_{i,t+1}|h_i)$ for dynasty i, we impose stationarity in (7) and recall that $u_{i,t+1}$ is independent from $v_{i,t+1}$ and $y_{i,t}$:

$$(1-\mu^2)\mathbf{Var}(y_{i,t+1}|h_i) = \mu^2\mathbf{Var}(v_{i,t+1}) + 2\mu^2\mathbf{Cov}(y_{i,t}, v_{i,t+1}|h_i) + \mu^2\mathbf{Var}(u_{i,t+1}) \tag{A.20}$$

To compute the covariance term, we use the stationarity of the process, the properties of $\epsilon_{i,t+1}$ and the properties of the covariance to take:

$$\mathbf{Cov}(y_{i,t}, v_{i,t+1}|h_i) = \frac{\rho_1\mu\sigma_v^2}{(1-\rho_1\mu)(1-\rho_1^2)} \tag{A.21}$$

Substituting (A.21) into (A.20), using the definitions of the variances for $v_{i,t+1}$ and $u_{i,t+1}$ and rearranging we obtain the expression given in the text, (16). The same reasoning yields the variance of talent for family i:

$$\mathbf{Var}(\theta_{i,t+1}|h_i) = \frac{1}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 \tag{A.22}$$

which is also increasing in μ. Taking the ratio of income's over talent's variance we obtain:

$$\frac{\mathbf{Var}(y_{i,t}|h_i)}{\mathbf{Var}(\theta_{i,t}|h_i)} = \frac{\kappa + \sigma_u^2}{\frac{\kappa}{\mu^2} + \sigma_u^2} \tag{A.23}$$

where we define:

$$\kappa(\mu,\rho_1) = \frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2}$$

Because $\mu < 1$, the denominator exceeds the numerator in (A.23), and the ratio is smaller than unity as claimed in Proposition 1. Next we prove the claim in Proposition 1 that this ratio is increasing in μ. The derivative of the ratio with respect to μ is proportional to:

$$\sigma_u^2\left[\kappa_1(1-\frac{1}{\mu^2}) + 2\frac{\kappa}{\mu^3}\right] + 2\frac{\kappa^2}{\mu^3} \tag{A.24}$$

where κ_1 is the derivative of κ with respect to μ. If the first term in (A.24) is positive, then our claim is proven. After some algebra, the sufficient condition reads as:

$$g(\mu,\rho_1) = \mu(\mu^2 - 1 - \mu\rho_1^2) > -1 \tag{A.25}$$

Because the function g has minimum at -1, ($\rho_1 = 1$ and $\mu = 1$), the sufficient condition holds and the claim is proven.

Finally, we consider the inequality in the cross section of families. From (16) it is obvious that $\mathbf{Var}(y_{i,t+1})$ increases in μ. For talent we have:

$$\mathbf{Var}(\theta_{i,t+1}) = \mathbf{Var}(\theta_{i,t+1}|h_i) + \frac{1}{(1-\mu)^2}\mathbf{Var}(\ln h_i) \tag{A.26}$$

where the first term in the right hand side of this equation is given by (A.22), and the last term equals the variance of the unconditional expectation of talent (the variance of (A.17)). It is straightforward to see that $\mathbf{Var}(\theta_{i,t+1})$ also increases in μ. From (16) and (A.26), consider the ratio of income over talent inequality in the cross section of families:

$$\frac{\mathbf{Var}(y_{i,t})}{\mathbf{Var}(\theta_{i,t})} = \frac{\kappa + \sigma_u^2 + \frac{1+\mu}{1-\mu}\mathbf{Var}(\ln h_i)}{\frac{\kappa}{\mu^2} + \sigma_u^2 + \frac{1}{\mu^2}\frac{1+\mu}{1-\mu}\mathbf{Var}(\ln h_i)} \tag{A.27}$$

where κ is defined above. To prove the claim in Proposition 1 that this ratio also increases in μ, let us define $\tau = \frac{1+\mu}{1-\mu}$, with $\tau' = \frac{2\tau}{1-\mu^2}$. Then after some tedious but straightforward algebra, the partial derivative of (A.27) with respect to μ is proportional to the following term:

$$\sigma_u^2\left[\kappa_1(1-\frac{1}{\mu^2}) + 2\frac{\kappa}{\mu^3}\right] + 2\frac{\kappa^2}{\mu^3} + \tau'\mathbf{Var}(\ln h_i)\sigma_u^2(1-\frac{1}{\mu^2}) + 2\frac{\tau}{\mu^3}\mathbf{Var}(\ln h_i)\left(\sigma_u^2 + 2\kappa + \tau\mathbf{Var}(\ln h_i)\right) \tag{A.28}$$

The first two terms of this expression are positive, as shown in (A.24) and (A.25). The term $2\tau\mathbf{Var}(\ln h_i)(2\kappa + \tau\mathbf{Var}(\ln h_i))/\mu^3$ is also positive. Therefore, after factoring out the term $\sigma_u^2\mathbf{Var}(\ln h_i)$, it suffices to show that:

$$\tau'(1-\frac{1}{\mu^2}) + 2\frac{\tau}{\mu^3} > 0 \tag{A.29}$$

Plugging in the definitions of τ and τ' and using the fact that $\mu < 1$, we can verify the above inequality.

A4. Intergenerational Correlation of Income and Talent

In this part we consider the intergenerational correlation within one dynasty i and treat h_i as a time invariant fixed effect. Because the variance is stationary, the stationary intergenerational correlation in income is equal to:

$$\textbf{Corr}(y_{i,t+1}, y_{i,t}|h_i) = \frac{\textbf{Cov}(y_{i,t+1}, y_{i,t}|h_i)}{\textbf{Var}(y_{i,t}|h_i)} = \mu + \mu\frac{\textbf{Cov}(y_{i,t}, v_{i,t+1}|h_i)}{\textbf{Var}(y_{i,t}|h_i)} \tag{A.30}$$

where we have used (7) and the properties of $u_{i,t+1}$. To obtain the expression (19) in the text, we insert in (A.30) the variance from (16) and the covariance in (A.21). We can differentiate (19):

$$\frac{\partial \textbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)}{\partial \mu} \propto \sigma_v^4(1-\rho_1^2)+\sigma_u^4(1-\rho_1^2)^2(1-\rho_1\mu)^2+\sigma_v^2\sigma_u^2(1-\rho_1^2)(2(1-\rho_1\mu)+\rho_1(1-\mu^2)) \tag{A.31}$$

Because all terms are positive, the correlation is increasing in μ and the claim in Proposition 1 is proven. A similar reasoning shows that the stationary intergenerational correlation of talent is:

$$\textbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i) = \frac{(\mu+\rho_1)\sigma_v^2 + \mu^3(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2}{(1+\rho_1\mu)\sigma_v^2 + \mu_t^2(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2} \tag{A.32}$$

Differently from income, the intergenerational correlation of talent has ambiguous comparative static in μ. A more progressive policy decreases both the covariance and the variance of income and talent. For income, the rate of decrease in the variance is smaller than that of the covariance and the comparative static is unambiguous. But in the case of talent, the covariance is not sufficiently decreasing because talent is not directly affected by μ. We can show that the intergenerational correlation in talent is increasing in μ provided that σ_u^2 is not too large relative to σ_v^2.

Finally, we prove the claim in Proposition 1 that the ratio of intergenerational correlations is smaller than one and increasing in μ. First, consider the ratio:

$$\textbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)/\textbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i) =$$

$$\frac{(\mu+\rho_1)(1+\rho_1\mu)\sigma_v^4 + \mu^3(1-\rho_1\mu)^2(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2(1-\rho_1\mu)(1-\rho_1^2)(\mu^2(\mu+\rho_1)+\mu+\mu^2\rho_1)}{(\mu+\rho_1)(1+\rho_1\mu)\sigma_v^4 + \mu^3(1-\rho_1\mu)^2(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2(1-\rho_1\mu)(1-\rho_1^2)(\mu^3(1+\rho_1\mu)+\mu+\rho_1)} \tag{A.33}$$

The difference between the last term in the denominator and the numerator is $\sigma_v^2\sigma_u^2(1-\rho_1^2)(1-\rho_1\mu)\rho_1(\mu-1)^2$. This difference is positive because $\sigma_v^2 > 0$, $\sigma_u^2 > 0$ and $\rho_1 < 1$. As a result, the expression in (A.33) is smaller than unity, strictly when $\mu < 1$, as claimed in Proposition 1. In addition, the ratio is increasing in μ. To see this, rewrite the ratio as:

$\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)/\mathbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i) =$

$$\frac{(\mu+\rho_1)\frac{1+\rho_1\mu}{(1-\rho_1\mu)^2}\sigma_v^4 + \mu^3(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}(\mu^2(\mu+\rho_1)+\mu+\mu^2\rho_1)}{(\mu+\rho_1)\frac{1+\rho_1\mu}{(1-\rho_1\mu)^2}\sigma_v^4 + \mu^3(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}(\mu^3(1+\rho_1\mu)+\mu+\rho_1)} \tag{A.34}$$

Denote by N the numerator and by D the denominator of this expression. The ratio of correlations increases in μ if and only if the derivate $N'D - D'N$ is positive. Since the denominator exceeds the numerator, $D > N$, it suffices to show that $N' > D' > 0$. From (A.34) it is evident that both terms increase in μ. From $N - D = -\sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}\rho_1(\mu-1)^2$, we take $N' - D' = -\sigma_v^2\sigma_u^2(1-\rho_1^2)\rho_1\frac{(\mu-1)(2-\rho\mu-\rho)}{(1-\rho_1\mu)^2} > 0$, which proves the claim.

A5. Proof of Proposition 2

Using the consumption function in equation (21) and the conditional expectation of income in equation (15), we can express the indirect utility function as:

$W(\mu_{t+1}; h_i, y_{i,t}, v_{i,t}) =$

$$= y_{i,t} + \ln\frac{\gamma}{\gamma+\mu_{t+1}} + \frac{1}{\gamma}\left(\alpha\ln\mu_{t+1} + \mu_{t+1}\left(\ln\left(\frac{\mu_{t+1}}{\gamma+\mu_{t+1}}\right) + \rho_0(1-\rho_1)\right)\right) + \frac{\mu_{t+1}}{\gamma}Q_{i,t} \tag{A.35}$$

where $Q_{i,t} = y_{i,t} + \rho_1 v_{i,t} + \ln h_i$ is family i's income potential at time t. Differentiating W with respect to μ_{t+1} we take:

$$\frac{\partial W}{\partial \mu_{t+1}} = W_1 + \frac{1}{\gamma}\left[W_2 + W_3 + W_4 + W_5 + Q_{i,t}\right] \tag{A.36}$$

In this expression, the term $W_1 = -\frac{1}{\mu_{t+1}+\gamma} < 0$ captures the intertemporal trade-off, $W_2 = \ln\left(\frac{\mu_{t+1}}{\mu_{t+1}+\gamma}\right) < 0$ measures the beneficial insurance effects of public policy, $W_3 = \frac{\gamma}{\mu_{t+1}+\gamma} > 0$ is the term associated with the distortions in investment, $W_4 = \frac{\alpha}{\mu_{t+1}} > 0$ is the direct output cost, $W_5 = \rho_0(1-\rho_1) > 0$ shows that insurance is less beneficial the higher is the long-run level of the endowment $v_{i,t}$, and $Q_{i,t}$ is defined above. Differentiating (A.36) with respect to μ_{t+1} we have:

$$\frac{\partial W^2}{\partial \mu_{t+1}^2} \propto \frac{1}{\gamma+\mu_{t+1}} - \frac{\alpha}{\mu_{t+1}\gamma} \tag{A.37}$$

A sufficient condition for single-peaked preferences is the strict concavity of the indirect utility. This requires that $\frac{\mu_{t+1}\gamma}{\mu_{t+1}+\gamma} < \alpha$. Since the left hand side of this inequality is bounded above by 1, the first part of the claim in Proposition 2 follows. For the second part of the Proposition, set $\partial W/\partial \mu_{t+1}$ equal to zero, and use the Implicit Function Theorem and the concavity of W in an interior optimum:

$$\frac{\partial \mu_{t+1}}{\partial Q_{i,t}} \propto \frac{\partial W^2}{\partial \mu_{t+1}\partial Q_{i,t}} = \frac{1}{\gamma} > 0 \tag{A.38}$$

A6. Proof of Proposition 3

If $0 < \mu_{i,t+1} < 1$ is the most preferred public policy for a dynasty with parameter $Q_{i,t}$, then it necessarily satisfies the first order condition, $\partial W/\partial\mu_{t+1} = 0$, where the derivative is given by (A.36). In addition, if $\alpha > 1$, then W is globally concave, and hence any solution to the first order condition will be the unique optimum. Since the Implicit Function Theorem applies, the comparative static $\partial\mu_{t+1}/\partial z$ has the same sign as the cross partial $\partial^2 W(\mu_{t+1}(h_i))/\partial(\mu_{t+1})\partial z$.

Therefore, $\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial\alpha} \propto 1/\mu_{i,t+1} > 0$, $\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial\rho_0} \propto 1-\rho_1 > 0$, $\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial\rho_1} \propto v_{i,t} - \rho_0$, and $\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial Q_{i,t}} = 1/\gamma > 0$. Since the most preferred policy μ_{t+1} of low $Q_{i,t}$ families is lower, it follows that when the position of the decisive agent p decreases, μ_{t+1} also decreases. For the parameter that expresses the degree of parental altruism, after some algebra and using the first order condition at optimum, we have:

$$\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial\gamma} = -\frac{1}{\gamma} < 0 \tag{A.39}$$

We briefly discuss the remaining comparative statics. First, social mobility is lower in societies with higher long-run income (higher ρ_0). At a first glance, this may appear counterfactual, since the conjecture is that in less developed economies, social mobility is lower (Solon, 2002). However, this could be because less developed economies have poorer tax collection technologies (high α) and limited expansion of voting rights (high p).

Second, in the original Becker and Tomes (1979) model, altruistic parents invest more in the human capital of their children which strengthens the intergenerational transmission and lowers social mobility. This result also holds in our model, but it takes place through a different mechanism.[26] Because a lower μ distorts fathers' investment decisions, a higher μ (less progressivity) redistributes resources in favor of the future generation. Hence, more altruistic fathers transfer more resources to the next generation by choosing a higher μ.

Third, if the decisive voter is temporarily well-endowed in family ability ($v_{i,t} > \rho_0$), then cultural persistence decreases the progressivity of the public policy. This result is consistent with the hypothesis that stronger family ties offer insurance and therefore "crowd out" the scope for social insurance.

Fourth, given the income potential $Q_{p,t}$, the parameters σ_v^2 and σ_u^2 do not affect the optimal μ. Because of the assumed log-log specification, substitution and income effects cancel off, and

[26] In our model, altruistic fathers invest more in their children human capital, holding constant μ. However, because of the log linear specification, altruism does not enter directly in the intergenerational transmission equation. See Solon (2004) for a similar result.

consumption and investment are constant fractions of output, independently of the properties of the shocks. In a more general specification of preferences, the scope for insurance will increase when endowment and market luck become more variable. Nevertheless, the properties of the two shocks can matter indirectly for μ, through the evolution of the income potential in the next period $Q_{p,t+1}$. Therefore, the persistence and volatility of the equilibrium μ are affected by cultural, genetic and market randomness.

A7. Proof of Proposition 4

First, we examine a stationary state with $\mu_{t+1} = \mu_t$. The population coefficient vector is defined as the argument that minimizes the least squares problem in the population:

$$(a, \beta) = \arg\min_{a,\beta} \mathbf{E}\left[(y_{i,t+1} - a - \beta y_{i,t})^2\right] \tag{A.40}$$

The well known formula for the population slope is given by:

$$\beta = \frac{\mathbf{Cov}(y_{i,t+1}, y_{i,t})}{\mathbf{Var}(y_{i,t})} = \mathbf{Corr}(y_{i,t+1}, y_{i,t}) = \frac{\mathbf{Cov}(\delta_0 + \mu_{t+1}\left(\ln h_i + y_{i,t} + v_{i,t+1} + u_{i,t+1}\right), y_{i,t})}{\mathbf{Var}(y_{i,t})} \tag{A.41}$$

which, from the imposed stationarity $\mathbf{Var}(y_{i,t+1}) = \mathbf{Var}(y_{i,t})$, also equals the cross sectional intergenerational correlation, $\mathbf{Corr}(y_{i,t+1}, y_{i,t})$. Recalling the properties of $u_{i,t+1}$ and $\epsilon_{i,t+1}$, we have:

$$\beta = \mu_{t+1}\left(1 + \frac{\mathbf{Cov}(v_{i,t+1}, y_{i,t}) + \mathbf{Cov}(\ln h_i, y_{i,t})}{\mathbf{Var}(y_{i,t})}\right) \tag{A.42}$$

The first covariance in the numerator is given by (A.21), because the fixed effect h_i is orthogonal to the $\epsilon_{i,t+1}$ and hence the $v_{i,t+1}$ process. The stationary covariance between the family fixed effect and income is given by:

$$\mathbf{Cov}(\ln h_i, y_{i,t}) = \frac{\mu_{t+1}}{1 - \mu_{t+1}}\mathbf{Var}(\ln h_i) \tag{A.43}$$

Putting all pieces together and setting $\mu_{t+1} = \mu_t = \mu$, yields the expression for β in Proposition 4.

Next, we show that β is increasing in μ. Using the variances in equations (16)-(18) yields:

$$\beta = \mu\left(\frac{\frac{\mu}{1-\mu^2}\frac{\mu+\rho_1}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 + \frac{1}{\mu}\mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i))}{\frac{\mu^2}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 + \mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i))}\right) \tag{A.44}$$

or

$$\beta = \frac{(\mu+\rho_1)\sigma_v^2 + \mu(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)}{(1+\rho_1\mu)\sigma_v^2 + (1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)} \tag{A.45}$$

Consider the last term in the numerator and the denominator. Because $\frac{1-\rho_1\mu}{1-\mu}$ is increasing in μ, this term also increases in μ. So, adding the same, increasing in μ, term both in the numerator and the denominator, tends, holding constant all other terms, to produce an increasing β, because the numerator is smaller than the denominator. Furthermore, β will increase more in μ due to this last term, when $\mathbf{Var}(\ln h_i)$ is higher. Hence, consider $\mathbf{Var}(\ln h_i) = 0$. In this case (A.45) collapses to the dynastic correlation in (19). Previously in this Appendix, we showed that this correlation is increasing in μ, which completes the proof of the claim that β increases in μ.

Differentiating (A.45) with respect to σ_u, we can show that:

$$\frac{\partial \beta}{\partial \sigma_u^2} \propto (\mu^2 - 1)\left(\rho_1\sigma_v^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)\right) \leq 0 \tag{A.46}$$

as claimed in Proposition 4. Differentiating (A.45) with respect to $\mathbf{Var}(\ln h_i)$, we obtain

$$\frac{\partial \beta}{\partial \mathbf{Var}(\ln h_i)} \propto (1-\mu^2)\left((1-\rho_1)\sigma_v^2 + (1-\rho_1^2)(1-\rho_1)\sigma_u^2\right) \geq 0 \tag{A.47}$$

The comparative statics of β with respect to α, p, $Q_{p,t}$, ρ_1 and γ follow from Proposition 3 and the result $\partial\beta/\partial\mu > 0$. Finally, we have verified numerically that μ is non monotonic in ρ_1 and σ_v^2 for various combinations of parameters.

Finally, for the second part of the Proposition we use the new equilibrium μ_{t+1} in the AR(1) process for income in (7). $\mathbf{Var}(y_{i,t})$ is given by (17) in the text for policy μ_t. The formulas for $\mathbf{Cov}(\ln h_i, y_{i,t})$ and $\mathbf{Cov}(v_{i,t+1}, y_{i,t})$ are taken by assuming that before the structural break the economy is in a steady state with $\mu_t = \mu_s$ for all $s < t+1$.

Appendix 2: Data

Social Mobility: Data for the intergenerational earnings elasticity is taken from Corak's (2006) meta-analysis. For Australia we use estimates from Leigh (2007). For Japan we use estimates from Lefranc, Ojima and Yoshida (2008). For Spain we use estimates from d'Addio (2007).

Private Return to Education: Taken from Boarini and Strauss (2007), Table 3. Calculated as the simple average in every country for the years available (males and females).

Total Government Spending and Social Welfare Spending: Government spending denotes central government consumption and investment. Social Welfare denotes consolidated government spending on social services as percentage of GDP. This data is taken from from Persson and Tabellini (2003). The variables are averaged over the 1960-1998 period.

Public Education: Data taken from OECD's Online Education Database. The series extracted are Public education expenditure as % of GDP, Public education expenditure per student (% of p.c. GDP), at all levels, and Public education expenditure per student (% of p.c.GDP), at the primary, secondary and tertiary level. For every country we average the series for all available years in periods 1970-2007.

Ethnolinguistic Fractionalization (ELF): Taken from Roeder (2001). The ELF index is defined as one minus the probability that two randomly chosen persons from a population belong to the same ethnic, linguistic or racial group. A higher ELF index denotes a more heterogeneous population. The value taken refers to the year 1961.

Gini Coefficient: The Gini coefficients at the factor and the gross earnings level are taken from Milanovic (2000) and are averaged across all available periods for any given country.

Weak Family Ties: Taken from Alesina and Giuliano (2007).

Political Inequality Variables: Taken from the Four Wave World Values Survey. The political participation variables that we use are recoded in binary form as follows: Interested in Politics (WVS code: $E023$; recoded as 1 for responders that answered 1 or 2, and 0 otherwise); Belong to Political Party ($A068$; already binary); Sign Petitions ($E025$; 1 if the responder answered yes and 0 otherwise); Participation in Lawful Demonstration ($E027$; 1 if the responder answered 1 or 2, 0 otherwise); Belong to Labor Union ($A067$; already binary). The income classification follows the variable $X047R$; see also Table 1.