The Political Economy of Intergenerational Income Mobility *

Andrea Ichino
UNIVERSITY OF BOLOGNA

Loukas Karabarbounis
HARVARD UNIVERSITY

Enrico Moretti
UNIVERSITY OF CALIFORNIA, BERKELEY

September, 2008

Abstract

The intergenerational elasticity of income is generally considered one of the best summary measures of the degree to which a society gives equal opportunity of success to all its members, irrespective of their family background. We present a parsimonious political economy model and show how the interaction between private and collective decisions determines the equilibrium level of mobility. Contrary to what it is generally assumed, a low correlation between father income and son income is not always desirable, as it may imply more inefficiency due to the distortionary effects of mobility-enhancing public policies. Moreover, taking into account the heterogeneity in preferences for intergenerational mobility leads to the conclusion that even if a fully mobile society is desirable ex ante, it may not be politically sustainable ex post. Our model clarifies the structural parameters behind the widely studied intergenerational elasticity of income in terms of political economy forces. Finally, we show some empirical evidence on the relationship between intergenerational elasticity of income across countries and its underlying determinants that is consistent with the predictions of the model.

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

*We would like to thank Giacomo Calzolari, Mattia Nardotto for insightful discussions and useful 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 opportunity 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* the intergenerational persistence of income. The widely studied intergenerational elasticity of income, β, is typically estimated with the Galton-Becker-Solon regression:

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

where y_s is son's income and y_f is father's income. A lower β denotes a smaller association between father's and son's income and therefore higher degree of social mobility.[1]

While much progress has been made on the empirical aspects of intergenerational mobility, less progress has been made on its economic interpretation. We still have only a partial understanding of what economic forces and institutions determine the parameter β and how they determine it. Important inroads in this direction have been made by Becker and Tomes (1979), who have showed how the intergenerational persistence of income reflects both "nature and nurture". In their model, individuals are given a certain 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 (1979) model generally ignores the role of redistributive institutions, and their deeper determinants. Such institutions have the potential to play an important role in determining how income is transmitted from one generation to the next. For example, the quality and fairness of the public education system can significantly affect economic opportunities of individuals who come from disadvantaged socioeconomic backgrounds. More in general, most redistributive policies—including affirmative action, welfare programs, subsidies that target children or poor parents—potentially affect the intergenerational elasticity of income. While some studies have highlighted the role of redistributive public policies as a determinant of social mobility, most existing studies take redistributive

[1] In a recent article, Solon (1999) provides a comprehensive survey of this literature.

policies as given.[2]

In this paper, we present a parsimonious political economy model with endogenous redistributive institutions. The model identifies the deeper economic and institutional determinants of social mobility when voters can choose redistributive institutions that consequently affect mobility. We show that the parameter β in the Galton-Becker-Solon regression depends on the genetic and cultural transmission of talent, parental investment in children and most importantly, the deeper determinants of redistributive policies.

The model focuses on how parents transfer economic endowment to their children through private and collective investment in their human capital. We assume that parents are altruistic, they know their own genetic ability but are ex ante 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 insurance against the risk of low genetic ability. We model these institutions as a redistributive scheme that takes away income from the better endowed children and gives it to the least endowed children.

An important insight of the model is that redistribution distorts parental private investment in children talent, and this generates a trade-off between insurance and incentives. Contrary to what most authors in the literature on intergenerational mobility assume, a low intergenerational elasticity of income (the parameter β) is not always necessarily desirable, as it may imply more inefficiency. In other words, it may be better for a society to have less intergenerational mobility in exchange for higher aggregate income. Moreover, in a world with heterogeneous dynasties, there is conflict of interests for the equilibrium level of social

[2]Examples of studies that have argued that institutions may be important determinants of mobility, and take institutions as given, include Glomm and Ravikumar (1992), Checchi, Ichino and Rustichini (1999), Solon (1999, 2004), Davies, Zhang and Zeng (2005), Mayer and Lopoo (2005), Hassler, Rodriguez Mora, and Zeira (2007), and Pekkarinen, Uusitalo and Kerr (2008). Two early papers that examine the determination of public policies in relation to social mobility are Benabou (1996) and Fernandez and Rogerson (1998). Bernasconi and Profeta (2007) also endogenize redistributive institutions and argue that the politically-determined level of public education may reveal the true talent of the children and therefore relax the mismatch of talents to occupations.

mobility. As a consequence, society will in general not choose the maximum amount of mobility, i.e. a zero intergenerational persistence of income ($\beta = 0$).

An additional important contribution of the model is that it clarifies how to interpret the widely studied intergenerational elasticity of income in terms of political economy forces. Specifically, we identify the structural parameters that underlie the parameter β in (1). We show how the equilibrium level of social mobility balances the costs of public education with its benefits. This trade-off is resolved by the political process that aggregates conflicting interests.

In the final part of the paper, we present some illustrative empirical evidence on the relationship between intergenerational elasticity of income across countries and its underlying determinants that is consistent with the predictions of our model. In a cross section of countries for which reliable estimates of β are available, we find a significant association between social mobility and the politico-economic variables that our our model predicts should affect mobility. For example, our model predicts that in countries where voter turnout is high and labor unions have strong bargaining power we should see stronger redistributive institutions like public schooling and consequently higher mobility.

On the other hand, in countries where the rich participate more in the political process than the poor, social spending for education and income mobility should be lower. We find that both these predictions are consistent with the data. Notably, the difference in the probability of political participation, as measured by the difference in the rate of party affiliation between rich and poor appears to be strongly correlated with β. Such difference has five times more predictive power than the private rate of return to education, which has been cited as one of the most prominent determinants of mobility by the literature. We also find that countries that are more ethnically and racially diverse and countries with stronger cultural transmission also tend to be less mobile. While causality is obviously unclear, these empirical associations are at least consistent with our model. In general, our politico-economic measures have more predictive power than previously studied determinants of social mobility.

The rest of the paper is organized as follows. In Section 2 we setup our model. Section 3 derives the politico-economic determinants of social spending and in Section 4 we show how these relate to the interpretation the standard intergenerational regression. In Section 5 we present our empirical evidence. Section 6 concludes.

2 A Simple Model of Intergenerational Transmission of Income

The main objective of our model is to derive the politico-economic parameters underlying the intergenerational persistence in income. This coefficient—β in equation (1)—has been the main focus of the existing empirical literature. Most of the theoretical work in this area has focused on the role of genetic transmission of ability, or the role of parental investment under different asset structures in explaining intergenerational transmission of income. Our framework builds on Becker and Tomes' (1979) model, as further explored by Goldberger (1989), Mulligan (1997) and Solon (2004).

In our setting, social mobility depends on redistributive policies that we model as outcomes 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 non human capital accumulation, it emphasizes the endogenous production of talent as an intermediate input for the production of final income.

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 intergenerational mobility. By endogenizing public policy, the well-known trade-off between incentives and insurance emerges naturally into the intergenerational mobility model. Piketty (2000) also makes this point.

2.1 Set-up

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 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 degree of redistribution; $\Theta_{i,t}$ is father's talent; and $U_{i,t}$ denotes a random and inelastic production factor 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. Redistribution 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 social policy, but also more 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 talented or lucky families, but more output for the less talented ones. One example of the social policy represented by μ_t is public education, since it increases output for low talented subjects but may distort output for the most talented ones.[3]

The parameter α characterizes the distortions associated with the redistributive system. For a given μ_t, a higher α implies that a smaller fraction of talents $\Theta_{i,t}$ gain from the redistributive system 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 redistribution. As α increases, the 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 redistribution.[4]

In each period t the following events take place:

[3]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 thus consistent with μ_t representing the progressivity of the public educational system in the production function (2).

[4]We do not restrict $\Theta_{i,t}$ to be smaller than unity. If, nevertheless, talent turns out to lie in the unit interval for all families in some period, then a public policy with $\alpha = 0$ can be thought as a growth enhancing reform that relaxes a credit constraint and promotes efficiently educational goals. The positive spillover effects of the policy would benefit every family, but the lowest talented families who face the binding constraint gain relatively more.

1. Fathers produce output $Y_{i,t}$, given talent $\Theta_{i,t}$ and the predetermined redistributive scheme μ_t.

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

3. Sons are born with a random endowment or ability $V_{i,t+1}$. 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.

We begin by treating μ as an exogenous parameter. In Section 3 we endogenize it. Son i is born with random *family endowment or ability* $V_{i,t+1}$. Following Becker and Tomes (1979), we assume that the logarithm of ability is 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 $\mathbf{E}(\epsilon_{i,t}) = 0$, $\mathbf{Var}(\epsilon_{i,t}) = \sigma_v^2$ and zero autocorrelations. We have $0 \leq \rho_1 < 1$ and therefore the logarithm of ability regresses towards the mean, $\mathbf{E}(v_{i,t}) = \rho_0$, and has stationary variance equal to $\mathbf{Var}(v_{i,t}) = \sigma_v^2/(1-\rho_1^2)$, where ρ_1 characterizes the cultural or genetic transmission 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 and therefore remains subject to the distortions introduced by the insurance scheme.

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.[5] $\gamma > 0$ parameterizes the degree of parental altruism, with higher values denoting smaller altruism. Parental investment $I_{i,t}$ can be thought as an intra-familiar, private educational input, for instance tuition fees, that increases a child's talent.

Sons' talent is produced according to:

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

where h_i is a family-specific time-invariant ability effect which allows dynasties to be ex ante heterogeneous. A higher h_i implies that family i has a higher long run ability level. We assume that h_i is distributed in the bounded interval $\mathbf{H} \subset R_{++}$ according to the density function Φ_h, and is orthogonal to the disturbances $\epsilon_{i,t+1}$ and $u_{i,t+1}$.[6]

2.2 Transmission of Income and Talent

In this Section we restrict attention to steady state redistributive schemes, i.e. $\mu_{t+1} = \mu_t = \mu$. 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 3, where we endogenize the redistributive policy.

Solving the problem in (4)-(5), using the production functions (2) and (6), and taking logs, we obtain the equation that describes for 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

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

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

[5]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).

[6]Becker and Tomes (1979) argue that ex ante heterogeneity is a realistic feature of the intergenerational mobility model as blacks in US and poorer families elsewhere tend to have lower long run incomes, for instance due to favoritism or market discrimination.

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

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

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

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

All the derivations in the paper are in Appendix 1.

The intercept $\delta_{0,i}$ can be decomposed into two parts. δ_0—which is a common effect across all dynasties i—and δ_i—which denotes the dynasty-specific time-invariant effect. Of course, families with higher h_i have higher lifetime income. Our δ_1 coefficient is different from the one in Becker and Tomes (1979) because we assume a multiplicative (in levels) production structure for output and talent. In our Cobb-Douglas environment, parental altruism γ does not affect the intergenerational transmission directly, i.e. for given policy μ (see also Solon, 2004, for a similar result).[7]

The slope δ_1 depends on the redistributive policy μ and is collectively decided by the fathers of each dynasty. In our model, social redistribution distorts the incentive of parents to invest in their children's human capital and, as a result, it weakens the intergenerational persistence of earnings. The link between deeper determinants of redistribution and intergenerational transmission of income is novel. While the previous literature has focused on the benefits of the private parental investment in children human capital, our model emphasizes the costs of public policies that seek to lower the intergenerational transmission of income.

In the Appendix we show that talent follows the stochastic difference equation

$$\theta_{i,t+1} = \lambda_{0,i} + \lambda_1 \theta_{i,t} + \lambda_2 v_{i,t+1} + \lambda_3 u_{i,t} \tag{14}$$

where

$$\lambda_{0,i} = \lambda_0 + \lambda_i \tag{15}$$

$$\lambda_0 = \ln\left(\frac{\mu}{\mu + \gamma}\right) + \alpha \ln \mu \tag{16}$$

$$\lambda_i = \ln h_i \tag{17}$$

$$\lambda_1 = \mu \tag{18}$$

[7]For simplicity, we normalize to one the exponent of investment in the talent production function (6) which represents the rate of return to parental investment.

$$\lambda_2 = 1 \tag{19}$$

$$\lambda_3 = \mu \tag{20}$$

The difference between the income and the talent transmission equations lies between the coefficients δ_2 and λ_2 (or δ_i and λ_i). These coefficients measure the output and talent elasticity of cultural or genetic ability. Because the redistributive system μ acts like a tax imposed on final output, the response of talent to genes is not affected by μ, and as a result talent can be thought as a more "private" production process. On the other hand, output's elasticity to family endowment does depend on the amount of redistribution because talent enters as intermediate input in the production of income, and income is subject to the redistributive instrument.

2.3 The Trade-Off Between Equity and Efficiency

2.3.1 Expectations

In our model, a more progressive and mobile system entails both costs and benefits for the society. To see this, consider the unconditional, stationary expectation or long run value of income for household i:

$$\mathbf{E}(y_{i,t+1}|h_i) = \frac{\mu\left[\rho_0 + \ln\left(h_i \frac{\mu}{\mu+\gamma}\right)\right] + \alpha \ln \mu}{1-\mu} \tag{21}$$

for all t. In (21), the expectation is conditioned on h_i to denote family dependency. There are four ways through which the redistributive system μ affects expected output.

1. *Distortions in Private Investment*: This is captured by the $\ln\left(\frac{\mu}{\mu+\gamma}\right)$ term. In more progressive systems (lower μ), the marginal propensity to invest in human capital, $\frac{\mu}{\mu+\gamma}$, 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 (2). The effects of redistribution on output are more adverse when the deadweight loss parameter α increases.

3. *Social Insurance or Benefits of Public Education*: The μ term that multiplies the bracket in the numerator of (21) captures the exponent of Θ^μ in the production function

(2). For low ability h_i dynasties, a more progressive social insurance or public education scheme increases expected lifetime income. The opposite happens for sufficiently able families, as shown in Figure 1.

4. *Intertemporal Insurance or Social Mobility*: This effect is given by the denominator $1-\mu$. For sufficiently high h_i dynasties, the numerator is positive and more mobility decreases expected income. On the other hand, low ability dynasties gain from the prospect of upwards mobility and progressivity increases their lifetime income.

Taking as given the realization of the state, father i's conditional expectation for son's income can be written as the sum of the long run level of income derived above and the transitory deviations of current income and endowment from their long run levels:

$$\mathbf{E}_t(y_{i,t+1}|h_i) = \mathbf{E}(y_{i,t+1}|h_i) + \mu\left(y_{i,t} - \mathbf{E}(y_{i,t+1}|h_i)\right) + \mu\rho_1\left(v_{i,t} - \rho_0\right) \tag{22}$$

where the time subscript in the left hand side denotes conditioning on the information set as of period t.

This analysis highlights two important points. First, there is a trade-off between equity and efficiency. This trade-off concerns solely the unconditional first moment of the income distribution: A lower steady state μ, causes permanent gains in terms of long run income to some families, while it hurts others. Second, there is a political conflict over the equilibrium level of social mobility. If families are ex ante heterogeneous, then fathers with permanently higher ability desire less progressive policies, everything else constant. Moreover, abstracting from long run differences due to h_i, fathers with transitory favorable realizations in their market activity, $u_{i,t}$, or family endowment, $v_{i,t}$, support a less progressive public policy.[8]

2.3.2 Variances

The stationary, unconditional variability that a given dynasty h_i faces in its income process is:

$$\mathbf{Var}(y_{i,t+1}|h_i) = \frac{\mu^2}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 \tag{23}$$

[8]Becker and Tomes (1979) and Solon (2004) consider a linear tax system and redistribution affects long run income only through channels (1) and (4). In this extension of their model, however, they treat public policy as exogenous. In our model, some families derive a net benefit from pro-mobility public policies, which makes the political economy aspect of the model interesting.

which 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 (23), we see that a more progressive redistributive system reduces overall variability in the income process. In addition, it lowers the fraction of variability attributed to endowment luck $v_{i,t+1}$. Intuitively, market luck $u_{i,t+1}$ affects only the final production of income, while family endowment $v_{i,t+1}$ matters both in the intermediate production but also through talent in the production of final output. As a result, a more progressive public education system will reduce the relative importance of the latter in the intergenerational variance of income. A weaker cultural or genetic transmission ρ_1 reduces variability in the income process, and similarly the fraction of endowment luck responsible for it.

If all families were identical, then the variance that families face in (23) coincides with the stationary ex post 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: [9]

$$\mathbf{Var}(y_{i,t+1}) = \mathbf{Var}(y_{i,t+1}|h_i) + \mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i)) \tag{24}$$

where the second term represents the variance "under the veil of ignorance", which from (21) is equal to:

$$\mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i)) = \frac{\mu^2 \mathbf{Var}(\ln h_i)}{(1-\mu)^2} \tag{25}$$

In (24), 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 redistributive system reduces all inequalities. The two components then differ in the role played by cultural or genetic persistence ρ_1. Because all families are assumed to transmit talent identically to their offsprings, ρ_1 does not matter for ex ante inequality, for given μ.

[9] In (23) the expression is not indexed by i and hence its expectation equals the expression itself. The dynastic variance is common across families i because we have assumed that h_i enters multiplicative and not exponentially into the production of talent. The same comment applies for the intergenerational correlation of incomes, in Section 2.3.3. Allowing for heterogeneity in the returns to investment would (i) not affect our theoretical claims, (ii) make the identification issue emphasized in Section 4 even more problematic, and (iii) complicate a potential empirical implementation of the model because the slopes of the regression—δ_1 and λ_1 in (7) and (14) —would depend on the observation i.

2.3.3 Covariances

Consider now the intergenerational correlation in incomes and talent. This summary statistic is what the literature calls social mobility, inequality across generations or "equality of opportunity". Conditioning on h_i, we distinguish between the intergenerational correlation of earnings within family, $\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)$, and the correlation we may observe in the data when families are heterogeneous, $\mathbf{Corr}(y_{i,t+1}, y_{i,t})$, and which is discussed later. Consider the time path of earnings and talent in 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,

$$\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i) = \frac{\mathbf{Cov}(y_{i,t+1}, y_{i,t}|h_i)}{\mathbf{Var}(y_{i,t}|h_i)} = \frac{(\mu + \rho_1)\sigma_v^2 + \mu(1 - \rho_1\mu)(1 - \rho_1^2)\sigma_u^2}{(1 + \rho_1\mu)\sigma_v^2 + (1 - \rho_1\mu)(1 - \rho_1^2)\sigma_u^2} \tag{26}$$

and that of talent, $\mathbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i)$, which is given in Appendix 1.

2.4 Summary

We summarize all the above findings in the following Proposition.

Proposition 1. Effects of Progressivity on First and Second Moments: *In any stationary state, with time invariant redistributive system $0 < \mu_{t+1} = \mu_t = \mu \leq 1$ we have:*

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

2. *The dynastic variance of income, $\mathbf{Var}(y_{i,t+1}|h_i)$, and that 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 strictly increasing in μ.*

3. *The cross sectional inequality of income $\mathbf{Var}(y_{i,t+1})$ and that 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 that a more progressive public policy decreases the dynastic variance in the production of income and talent, and decreases the cross sectional inequality. Moreover, in economies with more progressive redistribution, the within-dynasty intergenerational correlation of incomes is lower.[10]

In a society with no redistribution, the ratio of intergenerational correlations of income over talent and the ratio of the two variance components equal unity. This ratios decrease as public policy becomes more progressive. Intuitively, talent is an intermediate input and its intergenerational correlation weakens only because of the adverse effects of redistribution on parental investment. On the other hand, progressive public policy affects the intergenerational elasticity of income through two channels: Directly through decreased parental investment, but also indirectly by weakening the cultural or genetic transmission process that affects the production of income.[11] The bottom line for public policy is that redistribution may not be efficient in affecting directly the parental transmission of culture and genes, but it may still neutralize it indirectly because talent is an input in the production of income.

3 The Political Economy of Social Mobility

3.1 Politico-Economic Equilibrium

In this Section we endogenize the political process that aggregates conflicting preferences for intergenerational mobility. We assume that father i in period t observes the realization of last period's output, $y_{i,t}$, and endowment, $v_{i,t}$—which he takes as given—but does not observe the realization of his offspring's endowment $v_{i,t+1}$ and market luck $u_{i,t+1}$. We also assume that father has rational expectations about the realization of his offspring's endowment and market luck. His preference over the public policy μ_{t+1} is given by:

$$W(\mu_{t+1}; h_i, y_{i,t}, v_{i,t}, s) = \ln C_{i,t} + \frac{1}{\gamma}\mathbf{E}_t(y_{i,t+1}|h_i) \tag{27}$$

[10]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).

[11]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.

where s is the vector of structural parameters, and the conditional expectation, $\mathbf{E}_t(y_{i,t+1}|h_i)$, is given by (22). $C_{i,t}$ is the optimal level of consumption,

$$C_{i,t} = \frac{\gamma}{\mu_{t+1} + \gamma} Y_{i,t} \tag{28}$$

where we have reinstated the time subscript in the redistributive scheme.

A major simplification for our model is that sons are born after fathers have chosen the redistributive scheme μ_{t+1}, i.e. they do not affect its choice. Under this assumption, preferences of fathers over current policies are independent of future policies, and there is no need to explicitly consider dynasties' rational expectations about future policy outcomes.[12]

While this assumption is maintained mainly for analytical tractability, we believe it has intuitive appeal in the context of intergenerational mobility. As we show in Section 5 for a cross section of OECD countries, not every form of government activity is associated with social mobility. Rather, it is public spending on education that specifically matters for the strength of the intergenerational transmission. Since public education is regarded as highly redistributive at the primary level, i.e. before sons' political rights are extended, our setup captures this realistic feature of the intergenerational transmission.

The policy that maximizes (27) is

$$\mu_{i,t+1} = \mu(h_i, y_{i,t}, v_{i,t}; s) = \arg\max_{\mu} W(\mu; h_i, y_{i,t}, v_{i,t}, s) \tag{29}$$

where $s = (\alpha, \gamma, \rho_0, \rho_1, \sigma_u^2, \sigma_v^2)$ is the vector of structural parameters.

The most preferred redistributive scheme for every father trades off costs and benefits at the following levels. First, the four channels operating through the life long value of income analyzed in Section 2.3.1 apply. In addition, temporary deviations from the long run income value also affect the most preferred policy. Finally, redistribution allocates resources intertemporally. The consumption-investment ratio for every father is γ/μ_{t+1}.

[12]That is, the indirect utility W in (27) depends only on the current choice variable, μ_{t+1}, and not on future redistributive systems, $\mu_{t+2}, \ldots$. As a result, our model abstracts from 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 for the Markov Perfect Equilibrium in a non trivial dynamic voting game under the assumption of risk neutrality.

A less progressive system lowers investment and redistributes consumption in favor of the old generation. The opposite happens for a more progressive system. This intergenerational trade-off may capture a conflict in the allocation of total public expenditure between pensions directed to the old (higher μ) and public education that targets the young generation (lower μ).

We define total ability, $Q_{i,t}$, for family i at time t, as the sum of the life-long ability level $\ln h_i$, current income $y_{i,t}$ plus a term proportional to family endowment $v_{i,t}$:

$$Q_{i,t} = \ln h_i + y_{i,t} + \rho_1 v_{i,t} \tag{30}$$

Total ability, thus, is the sum of the life-long ability level plus the family-specific state variable. The latter summarizes the history of all relevant market and family shocks.

Proposition 2. Induced Preferences:

1. *Induced preferences over $\mu_{i,t+1}$ 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 $Q_{i,t}$.*

This Proposition establishes a sufficient condition for single peakedness. Second, it shows that families with higher life-long ability level h_i or temporarily favorable economic ($y_{i,t} > \mathbf{E}(y_{i,t+1}|h_i)$) or cultural ($v_{i,t} > \rho_0$) conditions prefer a less progressive public policy. This implies that the equilibrium redistributive system will not in general be time invariant, as assumed for simplicity in Section 2. There are various ways to proceed. The easiest case would be to assume a pre-commitment institution in which the initial fathers in period 0, observe $\{y_{i,0}, v_{i,0}, h_i\}$ and choose once and for all a time invariant redistributive scheme μ, which by assumption remains in operation for all future periods. A second possibility is to consider the steady state of the model, in which the distribution of abilities Q 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 settings, the analysis for the long run moments in Section 2 applies, with the time invariant coefficients for the income and the talent stochastic processes given by the optimal stationary μ.

A more realistic, and only slightly more complicated, environment would be to consider the transitional dynamics in a period-by-period decision making process. Under this setting, the equilibrium redistributive system (yet to be defined) will in general depend on the current state $Q_{i,t}$ of the decisive father. 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

$$\int_{Q_{l,t}}^{K} \phi_t(z)dz > 1/2 \tag{31}$$

where K is a constant, ϕ_t denotes the probability distribution of total ability in the cross section of dynasties as of the beginning of period t and $Q_{l,t}$ is the lowest realized total level of ability.[13] Under this setting, the expectations, variances and intergenerational correlations derived in Section 2 hold within each educational regime.

Letting $\Phi_t(Q) = \int_{Q_{l,t}}^{Q} \phi(z)dz$, we define the political institution in terms of the equilibrium outcome that it implies.

Definition 1. Institution P: *An institution P results in the redistributive scheme μ^e_{t+1} mostly preferred by the dynasty in the $100p$th percentile of the total ability distribution Φ_t, i.e. the family with a level of ability 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 and the political economy of taxation literature. Let the average level of total ability be $\bar{Q}_t = \int_{\mathbf{Q_t}} z d\Phi_t(z)$. Then if $p = \Phi_t(\bar{Q}_t)$, one obtains the utilitarian social rule that maximizes the welfare of the average father:

$$\max_{\mu} \int_{\mathbf{Q_t}} W(\mu, z; s) d\Phi_t(z) \tag{32}$$

[13] $K = \ln \frac{\gamma+\mu_l}{\gamma+\mu_h} \left(\frac{\gamma}{\mu_l-\mu_h} - 1 \right) - \ln \frac{\mu_h}{\mu_l} + \frac{\alpha}{\mu_l-\mu_h} \ln \frac{\mu_h}{\mu_l} - \rho_0(1-\rho_1)$. We index the distribution by t to show the possible dependency on μ_t and hence on calendar time.

Social preferences "behind the veil of ignorance", however, is not an entirely convincing aggregator of preferences, because in reality public policies are determined by the aggregation of *known* and conflicting political interests. The leading choice in the political economy literature is the one person, one vote pure majoritarian institution. If $\alpha > 1$, then by Proposition 2, induced preferences over policies $W(\mu, Q_{i,t}; s)$ are singe peaked. Therefore, the father with the median most preferred policy is the decisive voter. By the second part of the Proposition, this is the father with the median total ability level $Q_{50,t}$. The median father could be a high h_i ability person who faces unfavorable economic conditions in the market, or a father that belongs to a low ability family but faces a temporarily good shock. Note that this allows both the identity, and the income or the family endowment of the decisive father to change through 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). This also coincides with the unique equilibrium outcome in a representative democracy with two pure Downsian parties competing for fathers' vote.

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 (25). 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 the institution by p.

3.2 Politico-Economic Determinants

Given this definition, the properties of the equilibrium level of the public redistributive scheme, μ_{t+1}^e, are given in the following Proposition.

Proposition 3. Equilibrium Redistributive Scheme: μ_{t+1}^e *is increasing both in* α *and in* p*. It increases in* h_p*,* $y_{p,t}$*,* $v_{p,t}$ *and in* ρ_0*, decreases in* γ*, and does not depend on* σ_v^2 *and* σ_u^2*. It increases in* ρ_1 *if and only if* $v_{i,p} - \rho_0 > 0$*.*

This simple proposition shows that the redistributive system becomes less progressive (higher μ_{t+1}) when output costs α increase, but more progressive as the position p of the decisive dynasty in the ability distribution decreases. Our result shows that, as long as optimally chosen redistributive public policies affect intergenerational mobility, there is no reason to expect that a collective action of fathers transmits a perfectly mobile society to their sons. Note that for the refusal of this proposal, one would need to show *both* the costs associated with redistribution to be negligible *and* institutions to favor low ability families. This is an important point to keep in mind, because empirically it may difficult to find evidence for the magnitude of α or in reality some public reforms may entail small costs (Lindert, 2004). On the other hand, there is strong evidence that rich families have a larger "say" for the politico-economic outcome and the political system is wealth-biased (Benabou, 1996; Campante, 2007; Alesina and Giuliano, 2008).

We believe this is a point worth emphasizing because of a perceived gap between theory and practice in the literature.[14] The existing literature has attributed to the reduced-form coefficient in (1) a specific meaning. 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 perfect mobility case - $\beta_1 = 0$ - is often referred to as perfect 'equality of opportunity' because the income of a child is unrelated to the income of his or her parents. The degree of intergenerational mobility is therefore 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 analysis of Solon (1999), when he compares two countries and argues that A and B may have the same level of cross sectional inequality, but nevertheless their inequality be of very different character, because country A is perfectly mobile while B is perfectly immobile.

At the other extreme, Atkeson and Lucas (1992) have shown that when each dynasty consists of a single infinitely lived agent, the limiting distribution of utilities in a world with private information is degenerate and there is perfect inheritability of welfare and consumption. Recently, Phelan (2006) and Farhi and Werning (2007) challenge this result and argue that if dynasties are linked by a sequence of generations and the planner weights directly

[14] Piketty (2000) and Corak (2006) make a related point.

and positively all future descendants, then the social discount rate on future generations cannot exceed the private one. In this case, steady state efficient allocations involve a non degenerate distribution of utilities, mean reversion of consumption and positive amount of social mobility.

Relative to these papers, our model emphasizes—in addition to efficiency costs—political economy constraints that may further limit or enhance the extent of social mobility. For instance, perfect social mobility may be optimal in our model under a utilitarian institution (if α is very small), but not politically sustainable if rich families and special interests impose restrictions in the development of the welfare state and the provision of public education (if p is sufficiently high).

The politico-economic trade-off can be conceptualized by a decline in the position of the decisive voter p. Societies in which lower ability families have a larger "say" for the equilibrium outcome, choose more progressive redistributive systems, expect higher mobility and lower inequality. However, progressivity results in a lower lifetime level of income for sufficiently high ability families, and may even lower average income.[15]

In our model, the standard intuition, that the Rawlsian outcome entails a broader scope for government redistribution than the utilitarian optimum applies. Also, if the distribution of abilities Φ_Q is right skewed ($Q_{50} < \bar{Q}$)—perhaps because the ex ante ability distribution Φ_h is skewed—then a majority voting of fathers chooses a more progressive redistributive system relative to the utilitarian optimum. Holding average total ability $\bar{Q}$ constant, an increase in the (right) skewness of the distribution of abilities, leads the majority of fathers to demand more progressive policies and higher social mobility.

A higher ex ante inequality in abilities, $\mathbf{Var}(\ln h)$, could be associated with more skewness and hence a poorer median voter (lower Q_{50}). Alternatively, it could imply a smaller p if social preferences are averse to ex ante inequality. Therefore, we expect that higher ex ante inequality results in more progressive policies. On the other hand, if *de facto* political power is ultimately related to income, a higher ex ante variability could be associated with more powerful elites, lower redistribution and less social mobility. In Section 5 we offer some suggestive evidence that the second effect is more likely to dominate.

[15]In our model we do not explicitly consider 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 redistributive scheme increases long run income, implicitly capturing this realistic feature of public education.

The scope for beneficial reforms, and hence for social mobility, decreases 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, note that this is a ceteris-paribus proposition, and if in reality less developed economies are also less mobile, this is because of their poor technology in collecting taxes (high α) and the limited expansion of the voting rights (high p).

More altruistic parents (higher $1/\gamma$) choose less progressive public policies. 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. In the present model, altruistic parents also invest more in the human capital of their children. However, this *direct* effect of altruism on social mobility is absent from the intergenerational transmission because of the assumed Cobb-Douglas production structure and the log-log preferences. But our model highlights a second, novel, channel through which altruism affects the intergenerational transmission. On the one hand, altruistic fathers tend to insure their sons despite the distortion in their children's future human capital. But on the other, they sacrifice current consumption to increase the income of their children. In our setting the second effect turns out to dominate. Because a less progressive system redistributes consumption in favor of the young generation, altruistic parents transfer resources to the next generation by choosing a less redistributive scheme.

This simple example shows that in a more parameterized version of our model, deep parameters that the previous literature has shown to affect directly the intergenerational mechanism, will now also operate *indirectly* on social mobility, through the endogenously chosen public policy. These indirect effects may strengthen or offset the previously known direct effects.

Finally, if the decisive voter is temporarily well-endowed in family ability ($v_{i,t} > \rho_0$), then cultural persistence will decrease the progressivity of the redistributive system. This result is consistent with the hypothesis of Alesina and Giuliano (2007), that stronger family ties offer insurance and therefore "crowd out" the scope for social insurance.

As we discussed, the importance of the ex ante variance in (25) may be captured implicitly through the political process parameterized by p. On the other hand, σ_v^2 and σ_u^2 do not affect the optimal redistributive scheme, given the state of the system, $Q_{p,t}$. Because of the assumed log-log specification, substitution and income effects cancel off, and consumption and investment are constant fractions of output, independent of the shocks' properties. Obviously, with a more general specification of preferences, the scope for insurance would increase when endowment and market luck become more variable. Nevertheless, the properties of the two shocks can matter indirectly for the redistributive scheme, through the evolution of the state in the next period $Q_{p,t+1}$. Therefore, the persistence and volatility of the process for μ^e_{t+1} is affected by cultural, genetic and market randomness.

4 Interpretation of the Galton-Becker-Solon (GBS) Regression

The model with log-log utility and a double Cobb-Douglas production structure has the important advantage that it delivers the structural log-linear intergenerational earnings model which is the focus of the empirical literature.[16] The literature typically focuses on the Galton-Becker-Solon (GBS) regression:

$$y_{i,t+1} = a + \beta y_{i,t} + \varepsilon_i \tag{33}$$

where y_{t+1} and y_t denote son's and father's life long earnings in the population.

Previous models have recognized that β is a function of genetic and cultural inheritance, altruism and technological parameters, such as the net return to parental investment. However, we show that this coefficient also depends on institutions that a generation puts in place to insure the following generation from adverse shocks.

Proposition 4. Decomposition and Comparative Statics of Population Slope: *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.*

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

$$\beta_1 = \mathbf{Corr}(y_{i,t+1}, y_{i,t}) = \mu \left(1 + \frac{\frac{\rho_1 \mu}{(1-\rho_1^2)(1-\rho_1\mu)}\sigma_v^2 + \frac{\mu}{1-\mu}\mathbf{Var}(\ln h_i)}{\mathbf{Var}(y_{i,t};\mu)} \right)$$

   where the variance in the denominator refers to the cross sectional variance in (24).

[16] See Solon (1992), Zimmerman (1992), Bjorklund and Jantti (1997), Mulligan (1997), Solon (1999), and many others, for estimation of the log linear model. See Mazumder (2005, 2007), Lee and Solon (2006), and Aaronson and Mazumder (2008) for comparisons over time for the US.

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

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

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

According to the first part of the Proposition, in steady state, social mobility increases (β decreases) with the progressivity of the redistributive scheme (lower μ), the market luck variability (higher σ_u^2), and decreases with ex ante heterogeneity (higher $\mathbf{Var}(\ln h_i)$) for given policy μ. It decreases with output costs (higher α), the position and ability of the decisive family (higher p and Q_p), long run family endowment (higher ρ_0) and the degree of altruism (higher γ). The effects of cultural or genetical persistence ρ_1 and variability σ_v^2 are ambiguous.

The intergenerational elasticity of income β is defined as the covariance of incomes across generations divided by the cross sectional variance of income. The covariance between son's and father's income is smaller, the more progressive is public policy and the lower is the variability in market luck, cultural transmission and permanent ability in the population. These factors, however, also reduce the ex post variance in the cross section of families. In the case of μ_{t+1} and $\mathbf{Var}(\ln h_i)$, the covariance effect dominates, while with σ_u^2 the variance decreases faster. Intuitively, greater market variablity 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 redistribution. 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 enhance (if it is associated with smaller p) or decrease (under higher p) social mobility.

The second part of the Proposition sheds light on within country trends in the intergenerational elasticity of earnings. In particular, extension of the voting rights (lower p) or more efficient provision of public education (lower α) decrease the persistence of income across generations.

What do we learn from observing the population slope β? While β is interesting as a summary statistic of social mobility, unfortunately the GBS regression does not identify the deeper structural parameters.[17] Suppose that we can observe β for two countries A and B or two cohorts within a given country. In general, the condition for identification is that all but one parameter of interest be constant across countries or cohorts. For instance, with reference to the redistributive scheme μ, if inequality in the distribution of abilities, $\mathbf{Var}(\ln h)$, and cultural or genetic persistence, ρ_1, are identical across countries or cohorts, then from Proposition 4 there is a one-to-one correspondence between β and μ and μ is identified from the GBS regression. In general, this is a very restrictive assumption. Differences in the variance of the distribution of abilities or differences in cultural attitudes may drive cross country difference in β. For instance, Alesina and Giuliano (2007), report evidence on cross country differences in the organization of the family.

5 Empirical Evidence on the Politico-Economic Determinants of Mobility

The model in the previous Sections argues that politico-economic variables are important determinants of intergenerational mobility. In this Section we present some suggestive evidence that is consistent with the predictions of our model. Specifically, we seek to test whether pro-redistributive public attitudes and public policies are associated with social mobility, as predicted by our theoretical model.

An interesting study case is the experience of the UK over the past 50 years. Blanden, Gregg and Macmillan (2007) show that educational attainment accounts for a large fraction of the almost 50% decrease in social mobility between the 1958 and the 1970 cohort, measured at age 16. Our model suggests that a deeper determinant of the effect of educational inequality on social mobility may have been the sharp change in redistributive institutions that took place when Margaret Thatcher became Prime Minister in 1979. A sharp increase in regressive VAT taxes and a corresponding decrease in the more progressive corporate and income tax took place at the same time of a marked decline in the amount of public

[17] Goldberger (1989) first noted the under-identification of the Becker and Tomes (1979) model.

expenditure for education and a decline in the power of the unions.[18]

We now turn to cross country evidence. We stress that this exercise is purely suggestive. Two limitations preclude to draw strong conclusions. First, 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.[19] Second, the omitted variable problem is potentially severe since with such a small sample, we cannot control for all factors that affect social mobility.

In Figure 2, the vertical axis shows an estimate of β in a cross section of advanced democratic OECD countries. Consistent with the empirical literature on mobility, the UK, US and France appear as the least mobile societies, 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, 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.[20] 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. In Figure 3 we show that a similar correlation with mobility exists for total public expenditure in education as a percentage of GDP, public expenditure in education per student as a percentage of per capita GDP at the primary, secondary and tertiary level. Interestingly, the correlation is stronger when we measure education at the primary level, where expenditure is more redistributive.[21]

[18] 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.

[19] 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 also estimate the intergenerational income elasticity for Italy, but (i) the estimates found are especially high and (ii) even using a more conservative value, Italy is most of the times a major outlier. The only variable that seems to explain satisfactory Italy's low degree of mobility is the strength of family ties (high ρ_1). 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.

[20] 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).

[21] On the other hand, the correlation between intergenerational earnings elasticity and total government spending is only -0.05 and with spending on social expenditures around -0.11. The weakness of these

We now turn to more direct tests of the model. Specifically, we focus on empirical proxies for the politico-economic parameter p. Our first proxy for p is the turnout of the voters in national elections. If poor families are less likely to enfranchise, then our model indicates that in countries with a higher voter turnout there should be a more progressive public education system (smaller μ) and consequently more mobility (lower β).

In Figure 4 we show that there exists a positive correlation between turn out and public expenditures for education, and a negative correlation between turn out and β.[22] Of course, determining causality is difficult. For instance, Milligan, Moretti and Oreopoulos (2004) have shown that there is a causal effect of schooling on voting for the US.

Another prediction of the model is that in countries with high union density, the bargaining power of unionized fathers in shaping public policy should stronger and therefore p should be lower. The right portion of Figure 4 seems generally consistent with this prediction.[23]

To obtain a more direct measure of political preferences, we use data from the World Value Surveys dataset. We focus on the differential in political participation between poor and the rest of the population, where the rest of the population includes the middle class and the rich. The income classification into poor and non-poor follows the WVS and is standardized per country in our sample. On average, 34% of the population is classified as "poor" and the variation across countries is not large.[24] Political participation can be measured with a variety of variables. In Figure 5 we measure participation as probability of being member of a political party. The vertical axis is the fraction of the middle and the rich households that participate in political parties divided by the participation rate that prevails among the poor households. A lower value denotes a relatively more politically active class of poor families and hence a lower p. Note that we are not interested per se in the political participation of the poor, but in their participation *relative* to the other citizens of the same

relationships presents evidence that what may matter for social mobility is not in general the extent of government activity, but more specifically the educational expenditure, i.e. government activity directed at young ages. This is reasonable as the effects of educational expenditure are dynamic in nature; in contrast, we would expect social welfare spending on e.g. unemployment benefits, welfare programs, and assistance to poor families to be more closely associated to the degree of cross sectional inequality.

[22] Excluding the obvious outlier, US, does not overturn the signs of the associations, but decreases somewhat their magnitude. Specifically, the correlation with β falls from -0.56 to -0.35 and with public education it increases from 0.29 to 0.32.

[23] In the data, more leftist countries are associated with more public expenditure for education and higher mobility. The opposite holds for countries where the percentage of right votes and seats in the parliament is higher. However, these relationships are driven entirely by the US.

[24] The cross country range is 32%-35%, with Germany being the only outlier (39%).

country. Our measure of relative participation holds constant other factors that potentially drive political participation to be high in some countries and low in others, independently of income level.

The correlation in Figure 5 is consistent with the model.[25] If we compare the rich to poor gap in political participation, i.e. excluding the middle class, the correlation is slightly stronger (not shown). In Figures 6 and 7 we repeat 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 the one analyzed, with the bivariate correlations between our other proxies and β ranging from 0.43 to 0.63.[26] Quite strikingly, in a regression of β on one of the most commonly referred to determinants of social mobility, the private rate of return to human capital, the latter can explain only 8% of the cross country variation. On the other hand, the inequality between the rich and the poor families in participating in political parties explains 42% of the variation in social mobility.

In Figure 8 we present evidence that relates the degree of heterogeneity in a society to social mobility. Our model predicts that a higher ex ante heterogeneity (higher variance $\mathbf{Var}(\ln h_i)$) is associated with more public spending if p is low. But if p is high, then more heterogeneity is associated with higher ability for the decisive family, h_p, which in turn leads to a smaller degree of progressivity. Our proxy for the ex ante variance is the ethnolinguistic fragmentation index measured in 1961, which is defined as one minus the probability that two random persons in some country belong to the same ethnic, linguistic or racial group. The upper left panel shows that more diverse countries are associated with less public spending on education, which supports the view that the political system is wealth biased. The bottom panel shows that the direct link from ethnic, linguistic and racial diversity to less social mobility is also supported by the data. The bivariate correlation is 0.26, but excluding the very heterogeneous and mobile Canada, it 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 Alesina and Giuliano's (2007) index of weak family ties.[27] Weaker family

[25]The bivariate correlation of the latter with public spending is -0.49 and with the intergenerational elasticity 0.79. Excluding Germany the latter increases to 0.81.

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

[27]We thank the authors for kindly providing us with their data.

ties proxy for a lower ρ_1 in our model. In the right portion of Figure 8 panel, weaker family ties are associated with more public provision of education and more mobility.

We conclude with a final piece of evidence. Becker and Tomes (1979) original contribution aimed to explain within a unified economic model the degree of cross sectional inequality, and its relation with intergenerational inequality. We proxy for the cross sectional inequality in earnings, $\mathbf{Var}(y_{i,t})$, with the Gini coefficient measured at the *gross* earnings level. The variance in talent or skills, $\mathbf{Var}(\theta_{i,t})$, is proxied by the Gini coefficient measured at the *factor* income level. These statistics come from Milanovic's (2000) careful calculations.

Bjorklund and Jantii (1997) hypothesize that common causes may explain US's higher intergenerational and cross sectional inequality relative to Sweden's. Recently, Solon (2004) shows that cross sectional and intergenerational inequality are closely related, although they need not be perfectly correlated. The same result is implied in our model, where a more progressive educational system reduces both inequalities at the same time. 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. Figure 9 shows that the second force, which is more in line with our analysis, is more likely to be important. The bivariate association between the cross sectional gross earnings inequality and the intergenerational inequality as measured by β is around 0.72. Within our model, market variability, σ_u^2, explains the lack of perfect correlation as it increases cross sectional inequality to such a degree, where social mobility also increases.

Proposition 1 implies that the ratio of the gross earnings over factor inequality decreases in the progressivity of the educational system μ. In Figure 9, we see a strong association between the ratio of the Gini coefficients and public expenditure in education. In the same Figure, we show 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 theoretical model predicts that in societies where the poor participate more in political parties, labor unions and demonstrations, i.e. have a larger "say" for the equilibrium outcome, redistributive public education takes place and therefore the ratio of income over talent inequality decreases.

6 Conclusion

Intergenerational mobility is not randomly assigned to different societies, but emerges from the combined effect of "nature", "nurture" *and* public redistributive 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 highlights that redistributive policies may have distortionary effects. Taxation of the output of productive workers to transfer resources to less productive workers distorts the parental optimal private investment in children talent. This distortion generates a trade-off between insurance and incentives.

This implies that a low intergenerational correlation of income is not always necessarily desirable, because it comes at the cost of lower efficiency. Furthermore, in a world where parents have different endowment of talents, there is an inherent conflict of interests for the equilibrium level of social mobility. In our setting, redistribution is desirable to parents because it is an insurance against bad talent draws of their children. However, given a biological degree of intergenerational transmission of talent, better endowed parents prefer less redistribution than less endowed parents. As a consequence, the maximum amount of mobility is unlikely to be optimal.

Our empirical evidence on the relationship between intergenerational elasticity of income across countries and its underlying determinants appears generally consistent with the predictions of our model.

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. 2008. ”Preferences for Redistribution.” mimeo, Harvard University.

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.

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.

Bjorklund, Anders, and Markus Jantti. 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.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics,* 88(9-10): 1667–1695.

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.

Visser, Jelle. 2006. ”Union Membership Statistics for 24 OECD Countries.” *Monthly Labor Review,* January, 38—49.

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

Appendix 1: Derivations and Proofs

A1. Derivation of Income (7) and Talent (14) Transmission Equations

Solving the talent production function (6) for investment and then substituting into the resulting expression $\Theta_{i,t+1}$ from output's production function (2) for period $t+1$, we take:

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

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 $\mu_{t+1} < 1$. It then follows from our log-log specification of preferences in (4), that the solution is unique, interior and fully characterized by the first order condition:

$$\frac{C_{i,t}}{\gamma Y_{i,t+1}} = \frac{1}{\mu_{t+1}(h_i V_{i,t+1}) U_{i,t+1}} (\mu_{t+1})^{-\frac{\alpha}{\mu_{t+1}}} (Y_{i,t+1})^{\frac{1}{\mu_{t+1}}-1} \tag{A.2}$$

Solving for $C_{i,t}$ and substituting back to the budget constraint we derive the solution for children's income:

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

which when taking logs and letting $\mu_{t+1} = \mu$ yields the income transition equation (7) for the coefficients defined in (8)-(13). From (A.1) we get the solution for investment,

$$I_{i,t} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right) Y_{i,t} \tag{A.4}$$

which shows that investment equals a constant fraction of the endowment. Similarly, consumption is given by:

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

Finally, substituting the production function (2) into the solution (A.3), we derive the relationship between sons' income and talent of fathers:

$$Y_{i,t+1} = \left(\frac{\mu_{t+1}}{\mu_{t+1} + \gamma} \right)^{\mu_{t+1}} (h_i V_{i,t+1} U_{i,t+1})^{\mu_{t+1}} (\mu_{t+1})^{\alpha} \left[\mu_t^{\alpha} \Theta_{i,t}^{\mu_t} U_{i,t}^{\mu_t} \right]^{\mu_{t+1}} \tag{A.6}$$

Forwarding the output production function (2) one period and solving for talent, yields $\Theta_{i,t+1} = (\mu_{t+1})^{-\frac{\alpha}{\mu_{t+1}}} (U_{i,t+1})^{-1} (Y_{i,t+1})^{\frac{1}{\mu_{t+1}}}$. Substituting (A.6) into the latter and canceling terms we obtain the solution for talent:

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

Taking logs and setting $\mu_{t+1} = \mu_t = \mu$ gives the transmission equation for talent (14)-(20).

A2. Expected Income and Talent

First we show that given a redistribution parameter μ that is in steady state, 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 express the process for income in (7) as the sum of an ARMA(2,1) process plus an independent white noise:

$$y_{i,t+1} = (1-\rho_1)(\delta_{0,i} + \delta_2\rho_0) + (\delta_1 + \rho_1) y_{i,t} + (-\delta_1\rho_1) y_{i,t-1} + \delta_3 u_{i,t+1} - \delta_3\rho_1 u_{i,t} + \delta_2 \epsilon_{i,t+1} \quad (A.8)$$

The process is stationary if the roots of the characteristic equation

$$1 - (\delta_1 + \rho_1)x - (-\delta_1\rho_1)x^2 = 0 \quad (A.9)$$

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* or expectation or long run value of log income for family i is easily computed by setting $\mathbf{E}(y_{i,t+1}) = \mathbf{E}(y_{i,t}) = \mathbf{E}(y_{i,t-1})$ in (A.8) or (7), and the resulting manipulations yield equation (21) in the text. All comparative statics for the expectation analyzed in the text follow from inspection. A similar reasoning applied at the talent transmission equation (14) yields

$$\mathbf{E}(\theta_{i,t+1}|h_i) = \frac{\rho_0 + \ln\left(h_i \frac{\mu}{\mu+\gamma}\right) + \alpha \ln \mu}{1-\mu} \quad (A.10)$$

for all t.

From (7) and the definition of the unconditional expectation, the *conditional* on the state $\{y_{i,t}, \theta_{i,t}, v_{i,t}, u_{i,t}\}$ expectation of income is easily computed:

$$\mathbf{E}_t(y_{i,t+1}|h_i) = \mathbf{E}(y_{i,t+1}|h_i) + \mu\,(y_{i,t} - \mathbf{E}(y_{i,t+1}|h_i)) + \mu\rho_1\,(v_{i,t} - \rho_0) \quad (A.11)$$

where $\mathbf{E}(y_{i,t+1}|h_i)$ is the unconditional expectation given in (21). Similarly for talent we have:

$$\mathbf{E}_t(\theta_{i,t+1}|h_i) = \mathbf{E}(\theta_{i,t+1}|h_i) + \mu\,(\theta_{i,t} - \mathbf{E}(\theta_{i,t+1}|h_i)) + \rho_1\,(v_{i,t} - \rho_0) \quad (A.12)$$

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 $y_{i,t}$:

$$(1-\mu^2)\mathbf{Var}(y_{i,t+1}|h_i) = \mu^2\mathbf{Var}(v_{i,t+1}) + 2\mu^2\mathbf{Cov}(y_{i,t}, v_{i,t+1}|h_i) + \mu^2\mathbf{Var}(u_{i,t+1}) \tag{A.13}$$

For the covariance term, using the stationarity of the process, the properties of $\epsilon_{i,t+1}$ and that of the covariance we take:

$$\mathbf{Cov}(y_{i,t}, v_{i,t+1}|h_i) = \frac{\rho_1\mu\sigma_v^2}{(1-\rho_1\mu)(1-\rho_1^2)} \tag{A.14}$$

Substituting (A.14) into (A.13), 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, (23). We can follow the same reasoning and with minor modifications, the variance of talent for family i is given by

$$\mathbf{Var}(\theta_{i,t+1}|h_i) = \frac{1}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 \tag{A.15}$$

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

$$\frac{\mathbf{Var}(y_{i,t}|h_i)}{\mathbf{Var}(\theta_{i,t}|h_i)} = \frac{\kappa+\sigma_u^2}{\frac{\kappa}{\mu^2}+\sigma_u^2} \tag{A.16}$$

for $\kappa(\mu,\rho_1) = \frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2}$. If $\mu < 1$ and $\sigma_v^2 > 0$, then the denominator exceeds the numerator in (A.16), and the ratio is smaller than unity as claimed in Proposition 1.

To prove the claim in Proposition 1 that the ratio is increasing in μ, we can show that the derivative of the ratio with respect to μ is proportional to

$$\sigma_u^2\left[\kappa_1(1-\frac{1}{\mu^2}) + 2\frac{\kappa}{\mu^3}\right] + 2\frac{\kappa^2}{\mu^3} \tag{A.17}$$

where κ_1 is the derivative with respect to μ. Sufficient for the argument is that the first term is positive, or after some algebra that:

$$g(\mu,\rho_1) = \mu(\mu^2 - 1 - \mu\rho_1^2) > -1 \tag{A.18}$$

which proves the claim because the function g has minimum at -1, for $\rho_1 = 1$ and $\mu = 1$.

Finally, let's consider the inequality in the cross section of families. From (23) it is obvious that $\mathbf{Var}(y_{i,t+1})$ increases in μ. For talent we have

$$\mathbf{Var}(\theta_{i,t+1}) = \mathbf{Var}(\theta_{i,t+1}|h_i) + \frac{1}{(1-\mu)^2}\mathbf{Var}(\ln h_i) \tag{A.19}$$

where in the right hand side, the first term is given by (A.15) and the last term is the variance of the unconditional expectation of talent (where the latter is given in (A.10)). It is straightforward to see that $\mathbf{Var}(\theta_{i,t+1})$ also increases in μ.

The ratio of income over talent inequality in the cross section of families is therefore:

$$\frac{\mathbf{Var}(y_{i,t})}{\mathbf{Var}(\theta_{i,t})} = \frac{\kappa + \sigma_u^2 + \frac{1+\mu}{1-\mu}\mathbf{Var}(\ln h_i)}{\frac{\kappa}{\mu^2} + \sigma_u^2 + \frac{1}{\mu^2}\frac{1+\mu}{1-\mu}\mathbf{Var}(\ln h_i)} \tag{A.20}$$

To show that the ratio of variances 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.20) with respect to μ is proportional to the following term:

$$\sigma_u^2\left[\kappa_1(1-\frac{1}{\mu^2}) + 2\frac{\kappa}{\mu^3}\right] + 2\frac{\kappa^2}{\mu^3} + \tau'\mathbf{Var}(\ln h_i)\sigma_u^2(1-\frac{1}{\mu^2}) + 2\frac{\tau}{\mu^3}\mathbf{Var}(\ln h_i)\left(\sigma_u^2 + 2\kappa + \tau\mathbf{Var}(\ln h_i)\right) \tag{A.21}$$

By the argument laid out for (A.18), the first two terms are positive. Therefore, it suffices to show that

$$\tau'(1-\frac{1}{\mu^2}) + 2\frac{\tau}{\mu^3} > 0 \tag{A.22}$$

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

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. By the stationarity of the variance, the steady state intergenerational correlation in income is

$$\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i) = \frac{\mathbf{Cov}(y_{i,t+1}, y_{i,t}|h_i)}{\mathbf{Var}(y_{i,t}|h_i)} = \mu + \mu\frac{\mathbf{Cov}(y_{i,t}, v_{i,t+1}|h_i)}{\mathbf{Var}(y_{i,t}|h_i)} \tag{A.23}$$

where we have used (7) and the properties of $u_{i,t+1}$. Inserting the expression for the variance from (23) and the formula for the covariance in (A.14), and manipulating the resulting expression yields (26) as claimed. One can differentiate (26) and after some manipulations show that:

$$\frac{\partial\mathbf{Corr}(y_{i,t+1}, y_{i,t}|h_i)}{\partial\mu} \propto \sigma_v^4(1-\rho_1^2) + \sigma_u^4(1-\rho_1^2)^2(1-\rho_1\mu)^2 + \sigma_v^2\sigma_u^2(1-\rho_1^2)(2(1-\rho_1\mu)+\rho_1(1-\mu^2)) \tag{A.24}$$

All three terms are positive and hence, this proves the claim in Proposition 1.

A similar reasoning applies for the intergenerational correlation in talent, for which the stationary intergenerational correlation can be shown to be:

$$\mathbf{Corr}(\theta_{i,t+1}, \theta_{i,t}|h_i) = \frac{(\mu+\rho_1)\sigma_v^2 + \mu^3(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2}{(1+\rho_1\mu)\sigma_v^2 + \mu_t^2(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2} \tag{A.25}$$

which in general 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. On the other hand, the covariance of talent $\theta_{i,t}$ with ability $v_{i,t+1}$ is not sufficiently decreasing because talent is not directly affected by μ. The intergenerational correlation in talent is increasing in μ provided that σ_u^2 is not too large relative to σ_v^2.

To prove the claim in Proposition 1, we take ratio of intergenerational correlations by dividing (26) with (A.25). After some rearrangement we can show that this is:

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

$$\frac{(\mu+\rho_1)(1+\rho_1\mu)\sigma_v^4 + \mu^3(1-\rho_1\mu)^2(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2(1-\rho_1\mu)(1-\rho_1^2)(\mu^2(\mu+\rho_1)+\mu+\mu^2\rho_1)}{(\mu+\rho_1)(1+\rho_1\mu)\sigma_v^4 + \mu^3(1-\rho_1\mu)^2(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2(1-\rho_1\mu)(1-\rho_1^2)(\mu^3(1+\rho_1\mu)+\mu+\rho_1)} \tag{A.26}$$

The difference between the last term in the denominator and the numerator is $\sigma_v^2\sigma_u^2(1-\rho_1^2)(1-\rho_1\mu)\rho_1(\mu-1)^2$, which if $\sigma_v^2 > 0$, $\sigma_u^2 > 0$ and $\rho_1 < 1$, is positive. As a result, the expression in (A.26) is smaller than unity, strictly when $\mu < 1$, as claimed in Proposition 1.

Furthermore, the ratio is increasing in μ. To see this, rewrite the ratio as:

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

$$\frac{(\mu+\rho_1)\frac{1+\rho_1\mu}{(1-\rho_1\mu)^2}\sigma_v^4 + \mu^3(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}(\mu^2(\mu+\rho_1)+\mu+\mu^2\rho_1)}{(\mu+\rho_1)\frac{1+\rho_1\mu}{(1-\rho_1\mu)^2}\sigma_v^4 + \mu^3(1-\rho_1^2)^2\sigma_u^4 + \sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}(\mu^3(1+\rho_1\mu)+\mu+\rho_1)} \tag{A.27}$$

Denote by N the numerator and D the denominator. Then 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.27) it is evident that both terms increase in μ. The difference $N - D$ equals $-\sigma_v^2\sigma_u^2\frac{1-\rho_1^2}{1-\rho_1\mu}\rho_1(\mu-1)^2$, and therefore $N' - D' = -\sigma_v^2\sigma_u^2(1-\rho_1^2)\rho_1\frac{(\mu-1)(2-\rho\mu-\rho)}{(1-\rho_1\mu)^2} > 0$, which proves the claim in Proposition 1.

A5. Proof of Proposition 2

Using the conditional expectation of income defined in equation (22), the indirect utility can be expressed as:

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

$$= y_{i,t} + \ln \frac{\gamma}{\gamma + \mu_{t+1}} + \frac{1}{\gamma}\left(\alpha \ln \mu_{t+1} + \mu_{t+1}\left(\ln\left(\frac{\mu_{t+1}}{\gamma + \mu_{t+1}}\right) + \rho_0(1-\rho_1)\right)\right) + \frac{\mu_{t+1}}{\gamma} Q_{i,t} \quad \text{(A.28)}$$

where $Q_{i,t} = y_{i,t} + \rho_1 v_{i,t} + \ln h_i$ denotes total ability at time t for family i.

Differentiating W with respect to μ_{t+1} we take

$$\frac{\partial W}{\partial \mu_{t+1}} = W_1 + \frac{1}{\gamma}\left[W_2 + W_3 + W_4 + W_5 + Q_{i,t}\right] \quad \text{(A.29)}$$

where $W_1 = -\frac{1}{\mu_{t+1}+\gamma} < 0$ 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 as above.

Differentiating (A.29) once again with respect to μ_{t+1} we have:

$$\frac{\partial W^2}{\partial \mu_{t+1}^2} \propto \frac{1}{\gamma + \mu_{t+1}} - \frac{\alpha}{\mu_{t+1}\gamma} \quad \text{(A.30)}$$

A sufficient condition for single peakedness 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.

The second part of the claim follows straightforwardly by setting $\partial W/\partial \mu_{t+1}$ equal to zero, using the concavity of W in an interior optimum and finally taking:

$$\frac{\partial \mu_{t+1}}{\partial Q_{i,t}} \propto \frac{\partial W^2}{\partial \mu_{t+1}\partial Q_{i,t}} = \frac{1}{\gamma} > 0 \quad \text{(A.31)}$$

A6. Proof of Proposition 3

If $0 < \mu_{i,t+1} < 1$ is the most preferred redistributive system 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.29). 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 system 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 altruism, after straightforward algebra and using the first order condition at optimum, we have

$$\frac{\partial^2 W(\mu_{t+1})}{\partial\mu_{t+1}\partial\gamma} = -\frac{1}{\gamma} < 0 \tag{A.32}$$

as claimed.

A7. Proof of Proposition 4

First, assume we are in 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

$$(a, \beta) = \arg\min_{\beta} \mathbf{E}\left[(y_{i,t+1} - a - \beta y_{i,t})^2\right] \tag{A.33}$$

The well known formula for the population slope is given by

$$\beta = \frac{\mathbf{Cov}(y_{i,t+1}, y_{i,t})}{\mathbf{Var}(y_{i,t})} = \mathbf{Corr}(y_{i,t+1}, y_{i,t}) = \frac{\mathbf{Cov}(\delta_0 + \mu_{t+1}(\ln h_i + y_{i,t} + v_{i,t+1} + u_{i,t+1}), y_{i,t})}{\mathbf{Var}(y_{i,t})} \tag{A.34}$$

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:

$$\beta = \mu_{t+1}\left(1 + \frac{\mathbf{Cov}(v_{i,t+1}, y_{i,t}) + \mathbf{Cov}(\ln h_i, y_{i,t})}{\mathbf{Var}(y_{i,t})}\right) \tag{A.35}$$

The first covariance is still given by (A.14), 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:

$$\mathbf{Cov}(\ln h_i, y_{i,t}) = \frac{\mu_{t+1}}{1-\mu_{t+1}}\mathbf{Var}(\ln h_i) \tag{A.36}$$

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

To show that β is increasing in μ, we can use the expressions for the variances in (23)-(25), which yields

$$\beta_1 = \mu \left(\frac{\frac{\mu}{1-\mu^2}\frac{\mu+\rho_1}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 + \frac{1}{\mu}\mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i))}{\frac{\mu^2}{1-\mu^2}\frac{1+\rho_1\mu}{1-\rho_1\mu}\frac{\sigma_v^2}{1-\rho_1^2} + \frac{\mu^2}{1-\mu^2}\sigma_u^2 + \mathbf{Var}(\mathbf{E}(y_{i,t+1}|h_i))} \right) \tag{A.37}$$

or

$$\beta = \frac{(\mu+\rho_1)\sigma_v^2 + \mu(1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)}{(1+\rho_1\mu)\sigma_v^2 + (1-\rho_1\mu)(1-\rho_1^2)\sigma_u^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)} \tag{A.38}$$

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, given 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.38) collapses to the dynastic correlation in (26). 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.38) with respect to σ_u, we can show that

$$\frac{\partial\beta}{\partial\sigma_u^2} \propto (\mu^2-1)\left(\rho_1\sigma_v^2 + (1-\rho_1^2)(1+\mu)\frac{1-\rho_1\mu}{1-\mu}\mathbf{Var}(\ln h_i)\right) \leq 0 \tag{A.39}$$

as claimed in Proposition 4. Finally differentiating (A.38) with respect to $\mathbf{Var}(\ln h_i)$, we obtain

$$\frac{\partial\beta}{\partial\mathbf{Var}(\ln h_i)} \propto (1-\mu^2)\left((1-\rho_1)\sigma_v^2 + (1-\rho_1^2)(1-\rho_1)\sigma_u^2\right) \geq 0 \tag{A.40}$$

The comparative statics of β with respect to α, p, $Q_{p,t}$, ρ_1 and γ follow from Proposition 3 and the previously established $\partial\beta/\partial\mu > 0$. 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 steady state coefficient in the AR(1) process for income in (7), and $\mathbf{Var}(y_{i,t})$ is given by (24) in the text for policy μ_t.

Appendix 2: Data

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

Voter Turnout: Taken from IDEA. The voter turnout refers to all elections after 1945.

Union Density: Taken from Visser (2006), Table 3 and expressed as the percentage of actual relative to potential (based on eligibility) membership in unions. Averaged across four years: 1970, 1980, 1990 and 2000.

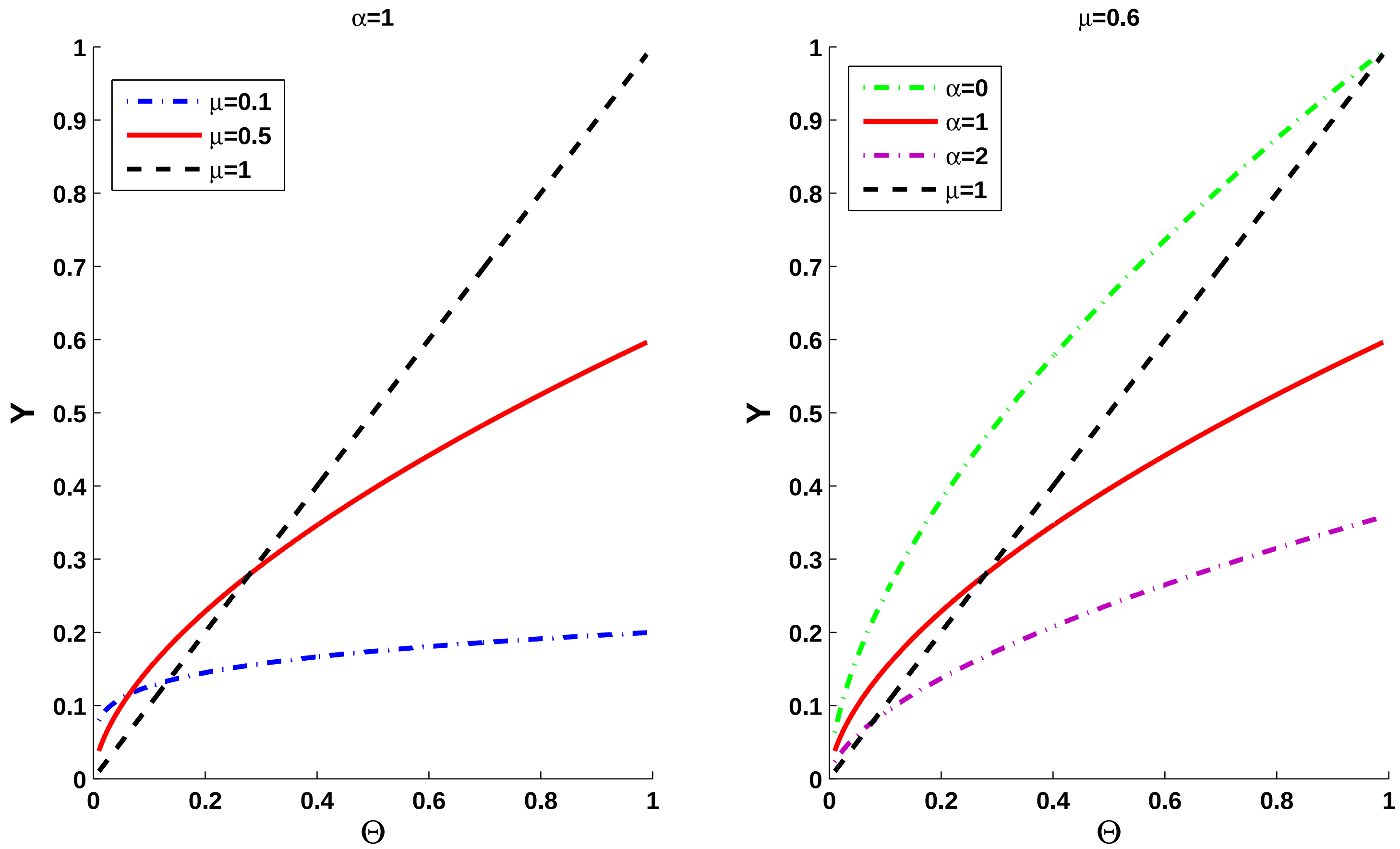
Gini Coefficient: The Gini coefficients at the factor and the gross earnings level are taken from Milanovic (2000) and are averaged across all available periods for any given country.

Weak Family Ties: Taken from Alesina and Giuliano (2007). A higher value denotes weaker family ties.

Left and Right Vote and Seats: The variables are taken from LIS-CWS database (2004) and averaged over all available years (1960-2000). Expressed as fractions of total votes / seats directed towards leftist and rightist parties respectively in the last elections. See LIS-CWS for the classification into leftish and rightist.

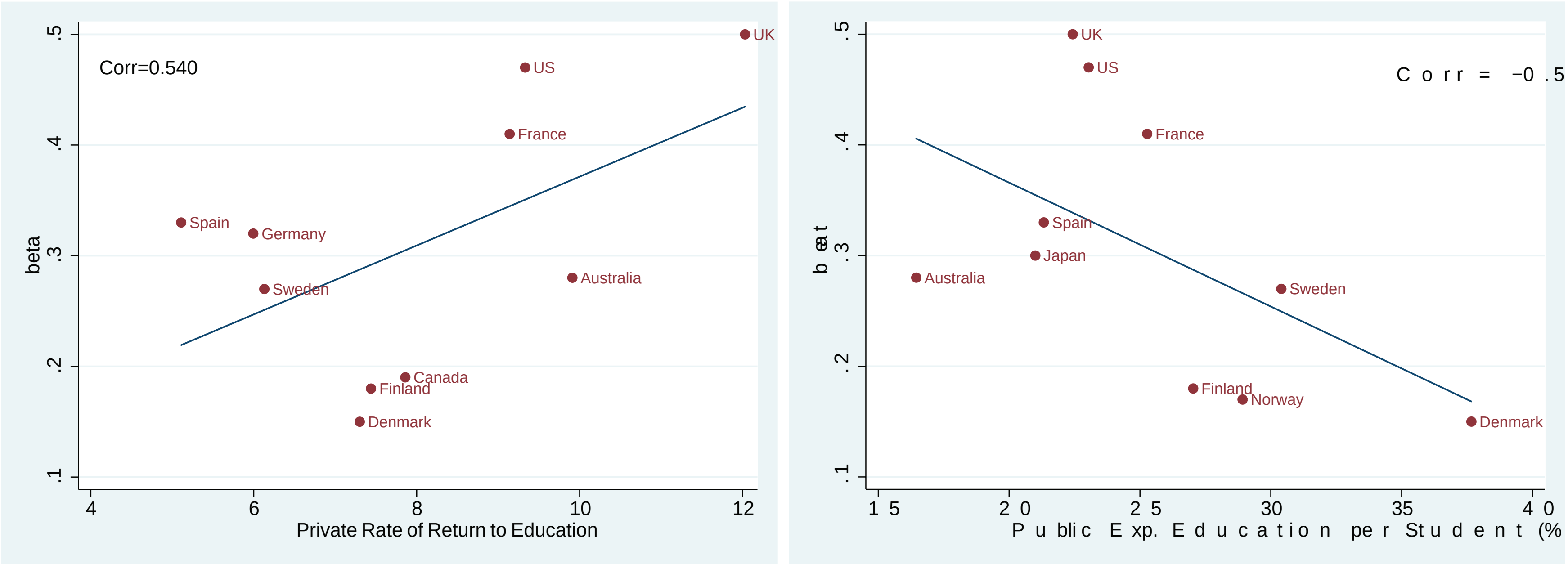
Political Inequality Variables: Taken from the 4-Wave World Values Survey. The political participation variables that we use are recoded in binary form as follows: Interested in Politics (WVS code: $E023$; recoded as 1 for responders that answered 1 or 2, and 0 otherwise); Belong to Political Party ($A068$; 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 $X047R$ that categorizes the responders into three categories, high, middle and low income. The classification is conducted at the national level and income is measured at the family level and is pre-tax and after transfers. The data is taken from the first observation (1981) or the 1990 if the former is not available.

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



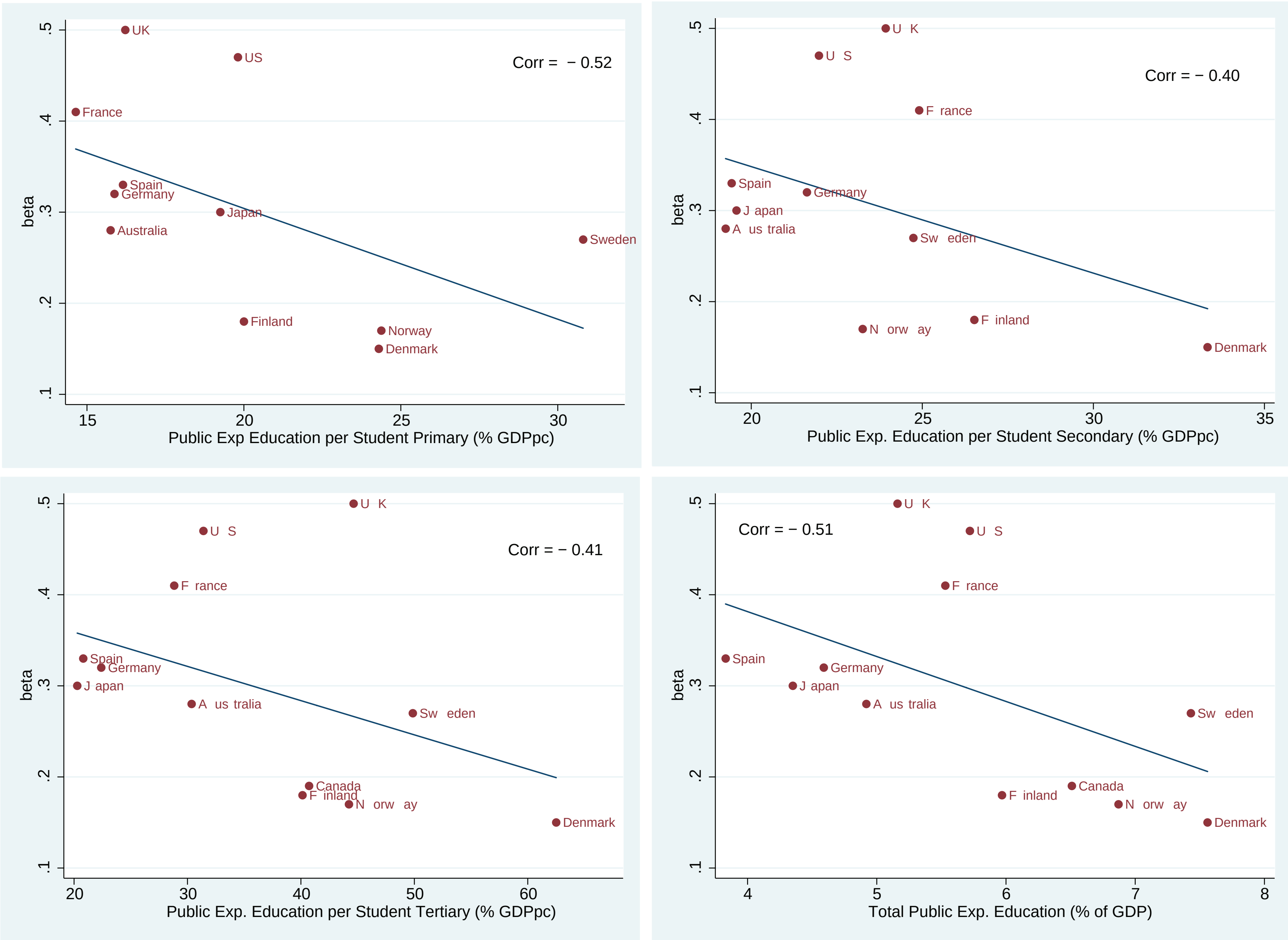
Note: For market luck normalized to $U_{i,t} = 1$.

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



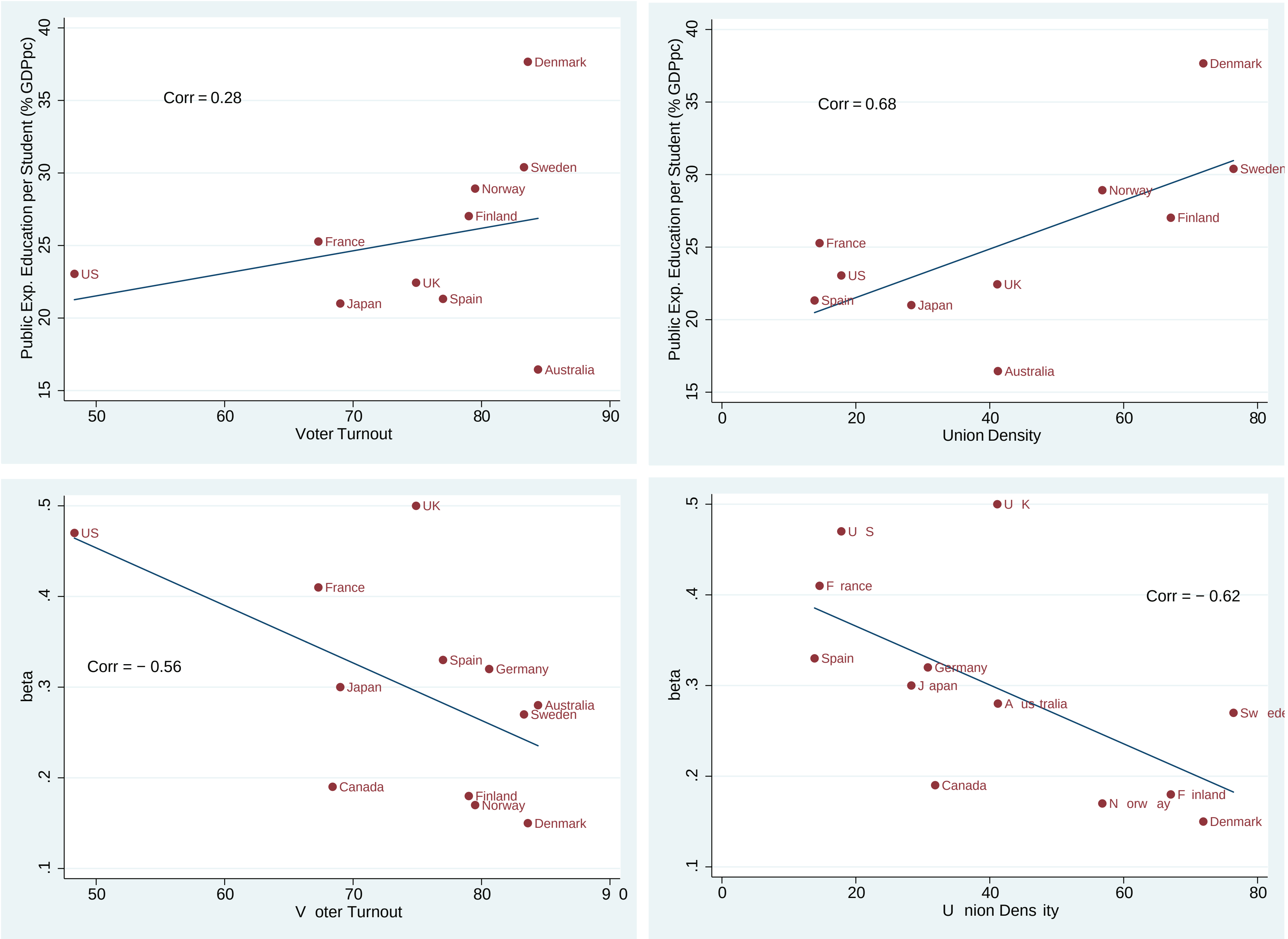
Note: The figure shows the bivariate relationship between the intergenerational earnings elasticity and the private rate of return to tertiary education (left panel) or the public expenditure in education per student as a percentage of per capita GDP. See Appendix for the data sources.

Figure 3: Other Measures of Public Education



Note: The figure shows the bivariate relationship between the intergenerational earnings elasticity and various measures of public education. See Appendix for the data sources.

Figure 4: Mobility, Public Education, Voter Turnout and Union Density

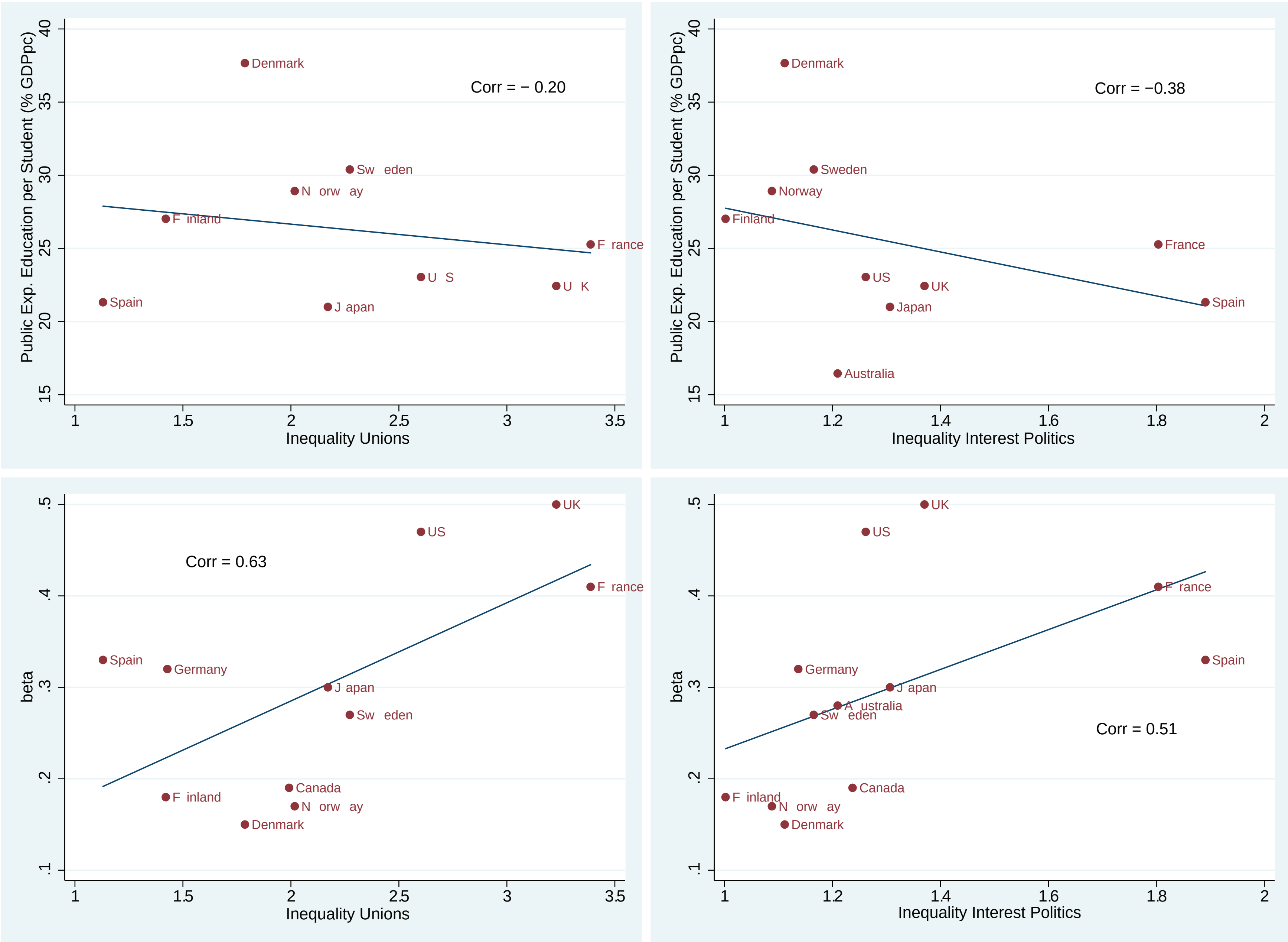


Note: See Appendix for the data sources.

Figure 5: Mobility, Public Education, and Inequality in Political Participation

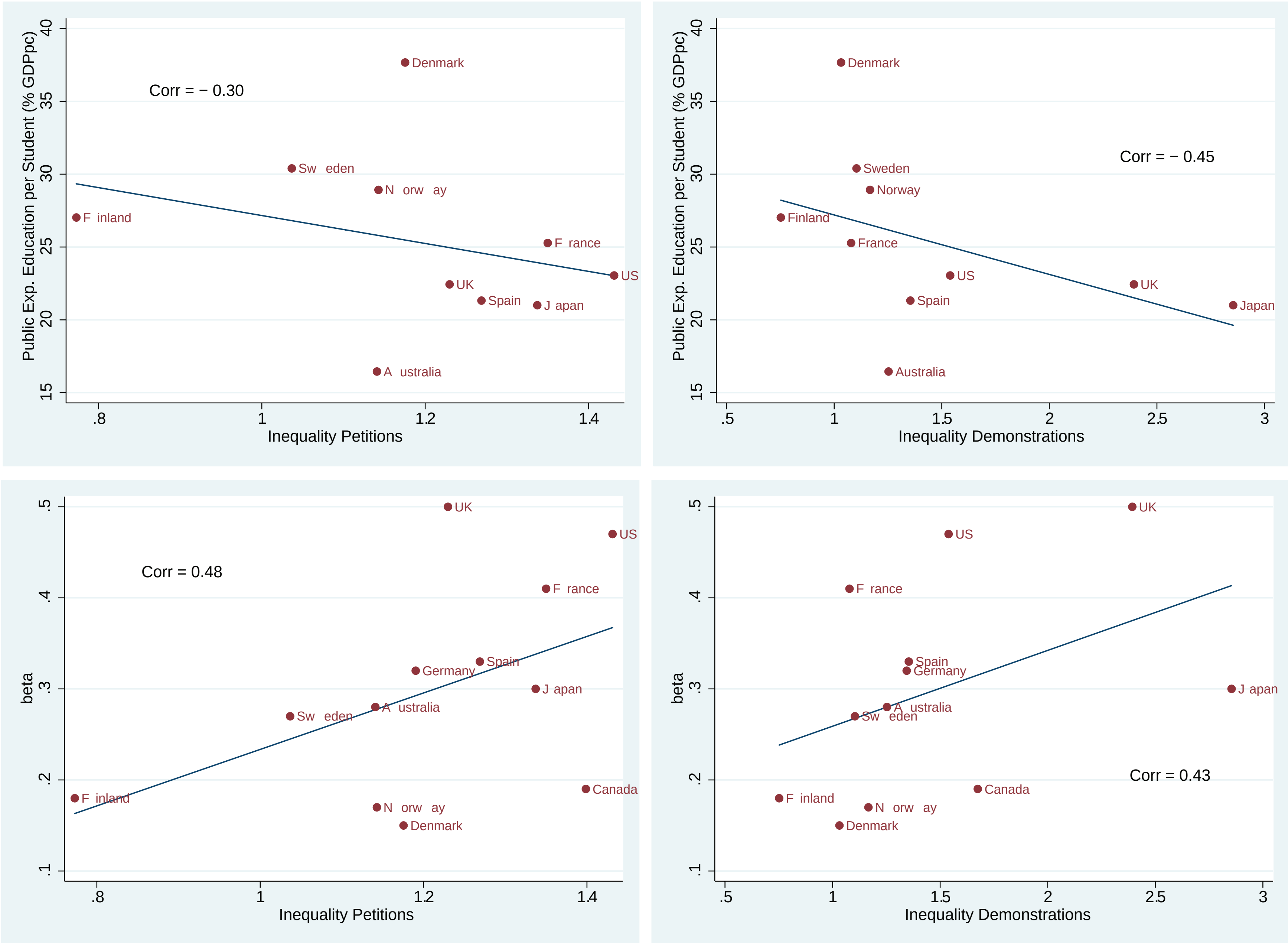
Note: Inequality Parties defined as the political party participation of the non-poor (middle and high income) citizens divided by that of the poor. See Appendix for the data sources.

Figure 6: Mobility, Public Education, and Inequality in Political Participation



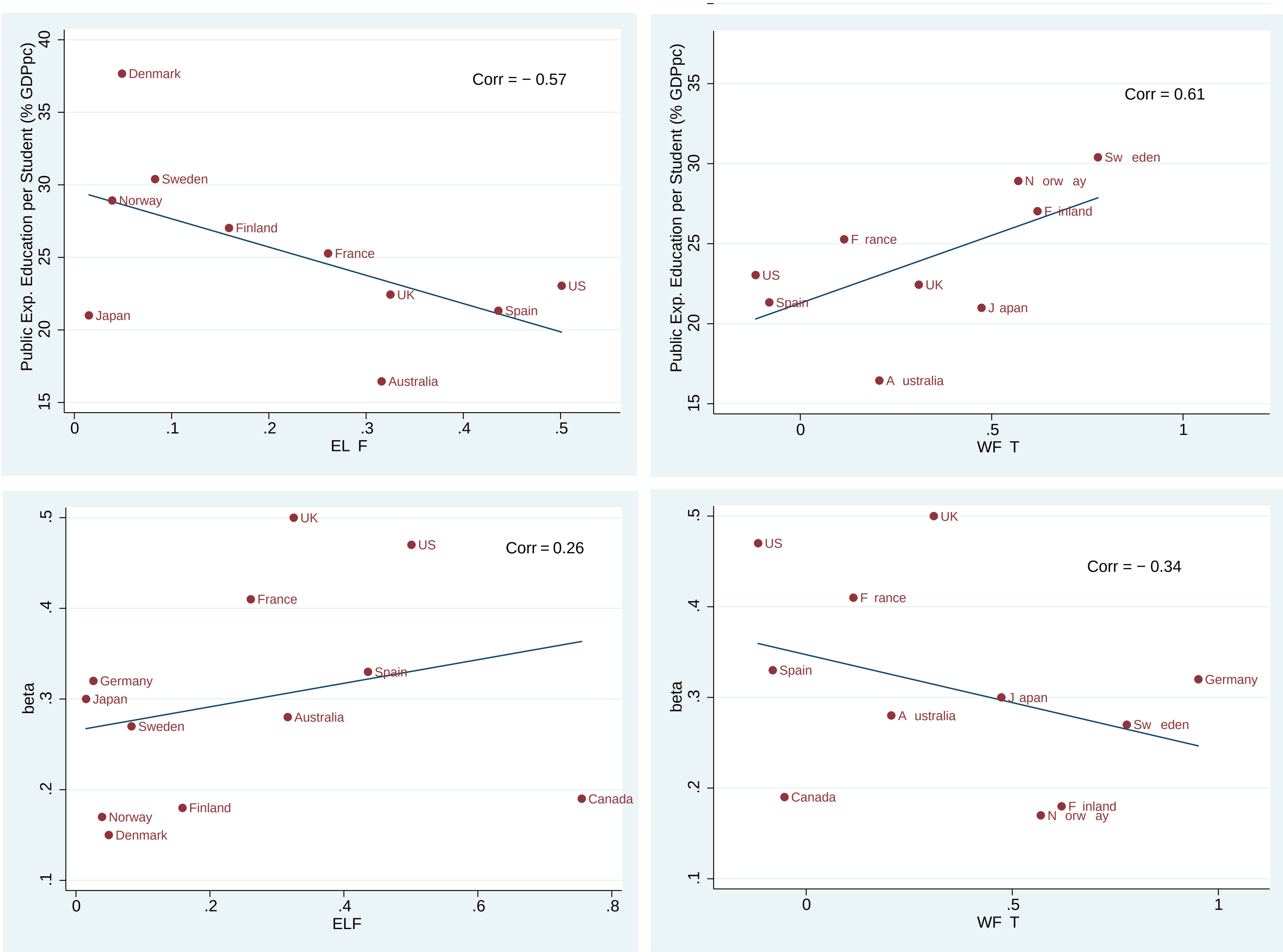
Note: Inequality Unions (Interest Politics) defined as the labor union participation (fraction answered "interested") of the non-poor citizens relative to that of the poor. See Appendix for the data sources.

Figure 7: Mobility, Public Education, and Inequality in Political Participation



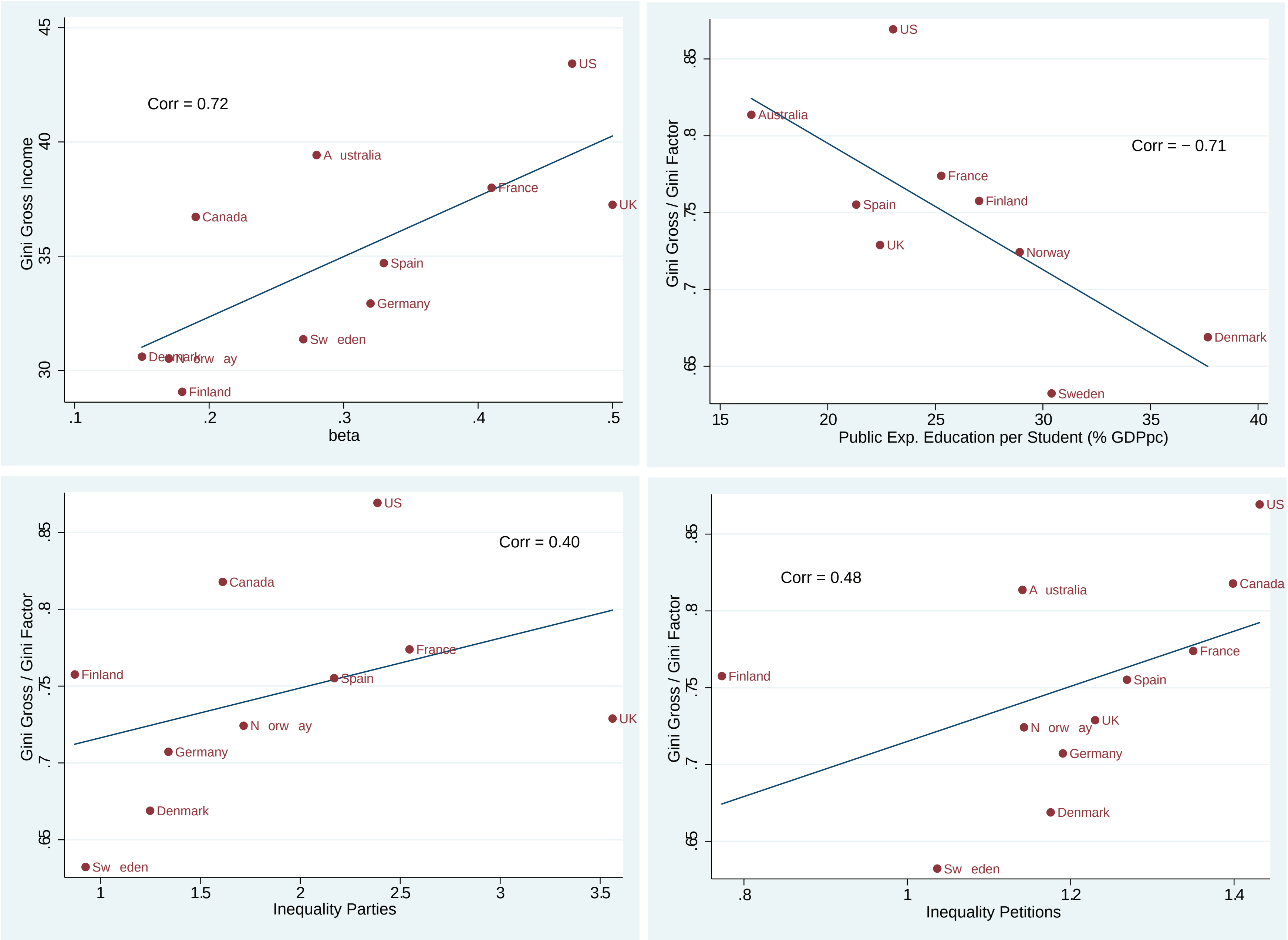
Note: Inequality Petitions / Demonstrations defined as the fraction of the non-poor citizens who sign petitions / participate in demonstrations divided by that of the poor. See Appendix for the data sources.

Figure 8: Mobility, Public Education, Heterogeneity and Family Ties



Note: The upper panel shows the bivariate relationships between public expenditure for education, the ethnolinguistic fractionalization (ELF) index and the weakness of the family ties (WFT). The bottom panel is the bivariate relationship of the intergenerational earnings elasticity with the ELF index and the index of WFT. See Appendix for the data sources.

Figure 9: Income and Talent Cross Sectional Inequality



Note: Inequality Parties / Demonstrations defined as the fraction of the non-poor (middle and high income) citizens who participate in parties / demonstrations divided by that of the poor. The gross and factor Gini coefficients are described in the text and the Appendix.