Human Capital Spill-Overs and the Geography of Intergenerational Mobility*

Brant Abbott

Queen's University

Giovanni Gallipoli

Vancouver School of Economics, UBC

June 15, 2016

Abstract

We develop and estimate an equilibrium model of geographic variation in the intergenerational elasticity of earnings (IGE). The theory extends the Becker-Tomes model, introducing a production sector in which workers' human capital inputs are complements. In this setting the return to parental human capital investments is lower where skill complementarity is more intense, and this is reflected in less intergenerational persistence. We also show that education subsidies may be more desirable where skill complementarities are stronger, endogenously leading to a negative correlation between progressive public policy and IGE. Using microdata we construct location-specific measures of skill complementarity and document that patterns of geographic variation in IGE are consistent with this hypothesis. Geographic differences in skill complementarity directly account for roughly one fifth of cross-country variation in IGE, and possibly more if one allows for the indirect effect through government expenditure in public education.

*We thank the editors and, in particular, Mariacristina De Nardi and Ronni Pavan for several comments and suggestions. The Russell Sage Foundation supported this research through a Presidential Authority Award. The Social Sciences and Humanities Research Council of Canada (SSHRC) funded Abbott through a postdoctoral fellowship while working on this research at Yale University. Gallipoli acknowledges SSHRC support through an Insight grant. We are grateful to Compute Canada and WestGrid for valuable computing power. Dimeng Chen, Colin Caines and Brad Hackinen provided excellent research assistance.

1 Introduction

A large literature documents significant differences in the intergenerational earnings elasticity (IGE) across countries (e.g. Corak, 2006; Black and Devereux, 2011; Jantti, Bratsberg, Røed, Raaum, Naylor, Osterbacka, Bjorklund, and Eriksson, 2006). Recent work has also measured intergenerational earnings mobility across regions of the United States (Chetty, Hendren, Kline, and Saez, 2014) finding large and persistent differences. Existing research documents interesting correlations between IGEs and various measures of public education spending, inequality and returns to human capital investments (see Blanden, 2009). Yet, less is known about what drives these correlations and whether they are useful for understanding differences in intergenerational mobility across regions.[1] An exception to this is the theoretical contribution by Ichino, Karabarbounis, and Moretti (2010), who show how variation in political institutions can lead to variation in both public education policies and IGEs.

In this paper we suggest that geographic variation in intergenerational mobility rates may be partly due to technology. We show that production complementarities in workers' human capital directly affect the intergenerational persistence of earnings. Differences in strategic complementarity may also result in differences in the desirability of progressive public education policies, highlighting an additional (indirect) channel through which technology may affect IGEs.

When we examine cross-sectional data about the geography of intergenerational persistence, we find evidence supporting this hypothesis. We use this reduced-form evidence to motivate our structural analysis and we estimate a richer model of earnings' persistence in which different countries ('islands') are characterized by different degrees of skill-complementarity in production. The model is parameterized using US and international data. Results suggest that differences in the degree of skill-complementarity can directly account for about 20% of international variation in IGEs. In comparison, observed variation in the generosity of public policies can explain about 25% of observed variation in IGEs.

Our theory of mobility highlights the importance of supply side factors. We start from the observation that each country's industrial composition spans several sectors, and workers within each sector have different skill endowments. Workers' skills are more or less substitutable depending on the sector. For example, workers' skills may be fairly complementary in the manufacture of complicated machinery, while in other industries, such as health or education, each worker's productivity is less dependent on the skills of co-workers.

[1]Quantitative studies of intergenerational persistence often focus on the 'aggregate' mobility rate within a country (e.g. Restuccia and Urrutia, 2004; Lee and Seshadri, 2014, for the United States).

To the extent that endowments, location and historical circumstances result in differences in the relative size of each industry within a country, one will observe heterogeneity in the level of skill substitutability across countries. Comparative advantage in certain industries may therefore influence human capital investments, government policies, and mobility between generations. Countries where industries employ technologies in which skills are more complementary will exhibit more mobility (i.e., less intergenerational income persistence) in equilibrium. Moreover, in these countries government policies that equalize skills would be more desirable. We present evidence of these relationships in cross-country data.

A key feature of our theory is that imperfect skill substitutability in production generates strategic complementarity in parental investments in children's human capital. The existence, and importance, of such strategic complementarity (or 'education spill-overs') in the United States has been documented by Moretti (2004)[2]. This means that the prevailing technology determines the degree to which a worker's own skills, as opposed to the skills of co-workers, determine her wage. In industries where skills are highly substitutable in production, a worker's wages are mostly determined by her own skills. Conversely, in industries where skills are relatively complementary in production, the skill levels of co-workers play a larger role through their effects on the overall productivity of the group.

This has direct consequences for the return to parental human capital investments. The more substitutable are skills, the greater the dependence of a worker's wage on her *own* skill attainment. Hence, the human capital investment made by parents will have a greater impact on their children's future earnings if they live in a country where skill substitutability is higher. Moreover, the greater returns to large human capital investments in countries where skill substitutability is higher will induce larger human capital investments among wealthy families.

Variation in the degree of skill substitutability may also have implications for the progressiveness of education and tax policies, and thus exert indirect effects on intergenerational mobility. As the degree of strategic complementarity in human capital investments increases, skills become more complementary and their homogeneity in the population induces significant improvements in the stock of human capital and aggregate productivity. A similar point has been made in the past by Arrow (1962) and Romer (1986). For this reason lower skill substitutability in production implies an increase in the desirability of policies that equalize skills, such as public education spending. Thus, the well-known association between progressive public policy and intergenerational mobility endogenously arises in such an environment. This observation highlights an additional channel through which skill substitutability may affect

[2]Moretti (2004) shows that spill-overs are larger between similar industries. His industry decomposition is finer, hence similar industries in his data mostly fall within the same coarse group at our level of aggregation.

intergenerational mobility.[3]

Finally, a negative association between income inequality and economic mobility arises in the model, as the degree of income inequality is directly related to the substitutability of skills in production. This is consistent with empirical observations suggesting that countries with more inequality also experience less earnings mobility across generations, a relationship that has been dubbed the ‘Great Gatsby’ curve (see Krueger, 2012; Corak, 2013).

We begin the paper with an analytical example, which illustrates the mechanism. In this simple setting we consider only two periods, and let parental human capital levels be exogenous endowments. There are no heritable skills, thus a parent can only affect her child’s outcomes by investing in her child’s human capital. Altruism motivates parents to do so; however, the skill substitutability parameter in the aggregate production function moderates the relationship between children’s future earnings and the human capital their parents bestow upon them. These simplifications lead to the stylized result that the IGE is proportional to the skill substitutability parameter. Furthermore, we show analytically that the optimal education subsidy and income taxation rates are decreasing in skill substitutability due to the lessening of strategic complementarity among skill investments.

The next step is to confront the implications of this simple theory with data. Our strategy is to develop measures of skill substitutability within industrial sectors using data for the US (O*Net survey data and CPS wage data). Given these measures we compute the average skill substitutability within each country, weighing industries by their relative size as measured in OECD-STAN data. Then we examine the cross-country relationship between aggregate skill substitutability and measures of the IGE taken from the literature. The results highlight a well-defined pattern: we consistently find a robust negative correlation between the IGE and the degree of skill substitutability prevailing in a country, as the analytical example suggests.

This empirical evidence indicates the presence of a relationship between industry composition and intergenerational income persistence. However, it does not establish the nature of this relationship. This motivates the next step, in which we develop a richer steady-state overlapping-generations model economy with an endogenous skill distribution and multiple sectors. Workers in this environment are free to sort into different industries, each with an associated level of skill substitutability. We characterize, and solve for, the equilibrium of the associated two-sided matching process.[4] The model is solved and estimated using US data. To assess the role of technological differences we run counterfactual experiments in which

[3]For a cross-country examination of the equilibrium effects of taxation on human capital accumulation, see Guvenen, Kuruscu, and Ozkan (2014).

[4]Total production of the final consumption good is a Cobb-Douglas aggregate of the goods produced in different sectors.

we change the industry composition to generate the degree of skill substitutability observed in other countries: this corresponds to re-weighting industries so that industrial composition replicates that of a different country where the intergenerational income elasticity is also known. Crucially, public education policy, marginal tax rates, and the progressiveness of the tax system are held constant at US levels in order to isolate the direct effect of technology. Our results indicate that almost one fifth of international variation in IGEs can be accounted for directly through differences in skill substitutability.

Lastly, we explore the possibility that human capital spill-overs may affect the IGE through the additional channel of optimal public policies. We solve constrained social planning problems for a select set of countries, and obtain the degree of skill subsidization that maximizes welfare conditional on the prevailing technology structure. When we allow for this indirect effect of technology on policy the model can account for up to one third of observed international variation in IGEs. A caveat about this exercise is in order: one should not draw conclusions about how much of the cross-country variation in IGE is due to differences in optimal policies, because the simple subsidy considered in the model often differs from the patchwork of observed public policies adopted by each country. Yet, we present evidence that countries characterized by lower levels of skill substitutability in production often adopt policies that encourage skill acquisition, fostering a more homogeneous work force.

The remainder of the paper proceeds as follows. Section 2 introduces a simple analytical model, develops the results and intuition relating skill substitutability to intergenerational mobility, and illustrates why the correlation between mobility and proxies of skill substitutability arise endogenously in such setting. Section 3 presents evidence of the empirical relationship between mobility and skill substitutability across countries. Section 4 describes the richer structural model and how we estimate it. Section 5 overviews the experiments and main findings, and Section 6 introduces an extension in which we allow for optimal skill subsidization, conditional on production functions. Section 7 provides a discussion of the results and concludes.

2 A Two-Generation Analytical Example

To illustrate the relationship between technology and the transmission of economic advantage we employ a two-period overlapping generations model. In the first period two generations are alive, adults and children. Adult agents are the parents of children who will become the adults of the second period. The setting can be generalized to include an infinite sequence of parents and children, but analytical tractability would be lost. In contrast, the two-period model yields clear and easily interpretable results.

2.1 Setup

A parent born into the initial generation is endowed with human capital, h_p. Parents derive income from their human capital, which allows them to pay for their own consumption, c_p, and for investments in their child's human capital, h_c. A child's human capital is equal to the investment made by the parent. When the child comes of age, she earns income based on her human capital and spends this income on own consumption, c_c.

Spill-overs are a key aspect of the production technology. High skilled workers exert a positive influence on each other's output, and on the productivity of lower skilled workers. In contrast, low skilled workers have a negative effect on the productivity of higher-skilled co-workers. One way to describe this type of interaction is to use a CES aggregator of the human capital supplied by all workers in the set I,

$$y = \left(\int_{i \in I} h_i^{\lambda} di \right)^{\frac{1}{\lambda}}. \tag{1}$$

This production function captures spill-over effects among co-workers, but it also has the uncomfortable implication that each worker is a *de facto* monopolist. To alleviate this downside we restrict human capital attainment to vary across discrete levels contained in the set H. Changing the variable of integration in the above formulation results in the following production function:

$$y = \left(\int_{h \in H} q(h) h^{\lambda} dh \right)^{\frac{1}{\lambda}}, \tag{2}$$

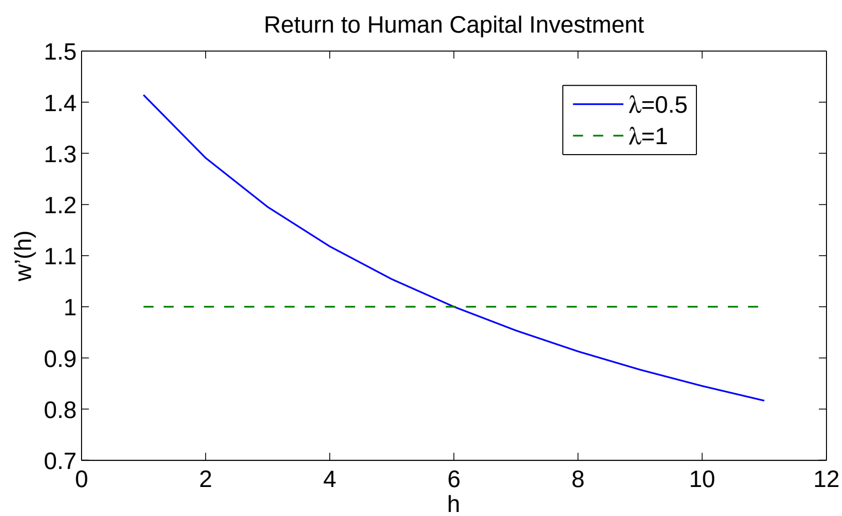
where $q(h)$ is the measure of workers from the set I who possess human capital h. Under this representation the representative firm chooses the measures $q(h)$ for each h so as to maximize profit, taking the wage function $w(h)$ as given.

The parameter $\lambda \in (0, 1]$ dictates the degree of skill substitutability. If $\lambda = 1$ then the skills of different individuals are perfectly substitutable. As λ becomes smaller the skills of different individuals become less substitutable. Lower substitutability of skills implies greater strategic complementarity in skill investments because the productivity of each worker increasingly depends on the productivity of co-workers. This is apparent in the competitive wage of a worker with human capital h, which equals the marginal product of the measure $q(h)$:

$$w(h) = \frac{1}{\lambda} y^{1-\lambda} h^{\lambda}. \tag{3}$$

Other workers' productivities influence a person's wages through y, but the influence of y diminishes as λ approaches unity. An implication of this wage function, and a key difference from the standard Becker and Tomes model, is that the marginal return to human capital investments is positive but decreasing in h. That is, $w'(h) = y^{1-\lambda} h^{\lambda-1}$. This follows because, in the presence of skill complementarity, the return to human capital investments is also a function of the 'fixed factor' y. Figure (1) illustrates this point, plotting the marginal return to human capital investments for the linear ($\lambda = 1$) case and for a high-complementarity case ($\lambda = 1/2$).

Figure 1: Marginal return to human capital investments under different levels of skill complementarity.



The parent's value function, $V_p(h_p)$, depends on own consumption and, through altruism, on the utility of the child, $V_c(h_c)$. The degree of altruism is equal to β, hence the parent's value function is

$$V_p(h_p) = \max_{c_p, h_c} \left\{ u(c_p) + \beta V_c(h_c) \mid c_p + h_c = w(h_p) \right\}. \tag{4}$$

The constraint in the maximization problem is the parent's budget identity. Own consumption and investment in the child must be paid for out of labor income $w(h_p)$. Implicit in this restriction is that parents cannot borrow against the child's future income in order to finance a child's human capital.

2.2 Implications for Earnings Mobility

Given the two period nature of the problem the child's value function is simply,

$$V_c(h_c) = \max_{c_c} \{u(c_c) \mid c_c = w(h_c)\}. \tag{5}$$

The constraint is the child's budget identity. Because the child is the final generation she simply consumes all earnings.

The following inter-generational equation can be derived from the first-order optimality conditions:

$$h_c^{1-\lambda} = \frac{u'(c_c)}{u'(c_p)} \beta \lambda y_c^{1-\lambda}. \tag{6}$$

This equation obviates the complexity of the parent's problem. The choice of human capital investment in a child is related to the marginal utility of consumption of both child and parent, as well as to aggregate output.

Under the assumption of log-utility a child's human capital can be expressed as a function of primitive parameters:

$$h_c = \frac{\beta\lambda}{1+\beta\lambda} y_p^{1-\lambda} h_p^{\lambda}. \tag{7}$$

This expression has the stark implication that the elasticity of a child's earnings with respect to parental earnings is directly dependent on the prevailing degree of skill substitutability in the economy. A second important implication is that the influence of other people's skills on one's own human capital increases as skills become more complementary in production.

The same expression can also be written in terms of income. Because $w(h)$ is earnings, equation (7) can be rearranged into an intergenerational earnings equation:

$$w(h_c) = \left(\frac{\beta\lambda}{1+\beta\lambda}\right)^{\lambda} y_c^{1-\lambda} w(h_p)^{\lambda}. \tag{8}$$

Thus, the intergenerational elasticity of income also depends directly on λ.

2.3 Optimal Education Policy

In cross-country data the level of public support appears negatively related to the intergenerational income elasticity, indicating that larger public education subsidies are associated with greater economic mobility. What remains unclear is whether there is a causal relationship, or whether this coincidence is the by-product of mutual associations with some other variable.

In this simple analytical example the optimal subsidization of human capital investments (in the Ramsey sense) depends on skill substitutability in production. Less substitutability implies that subsidization is socially more desirable. Hence, the optimal level of public education support is negatively related to λ, and positively related to intergenerational mobility.

Next, we study a constrained social planning problem in which ex-ante social welfare is maximized by choosing a proportional subsidy for human capital investments, s, as well as a proportional wage tax, τ. Ex-ante welfare is the expected discounted utility of a family behind the 'veil of ignorance', prior to learning the relative advantage of the parent. Hence, the planning problem is:

$$\max_{s,\tau} \int V_p(h_p; s, \tau) dF(h_p)$$

$$s.t.$$

$$\begin{aligned} c_p + (1-s)h_c &= (1-\tau) y_p^{1-\lambda} h_p^{\lambda} \\ c_c &= y_c^{1-\lambda} h_c^{\lambda} \\ h_c &= \tfrac{1-\tau}{1-s} \tfrac{\beta\lambda}{1+\beta\lambda} y_p^{1-\lambda} h_p^{\lambda} \\ \tau y_p &= s \textstyle\int h_c di \\ y_c &= (\textstyle\int {h_c}^{\lambda} di)^{\frac{1}{\lambda}} \end{aligned} \tag{9}$$

Crucially, the third constraint imposes that household optimization holds, making this a Ramsey planning problem. The fourth constraint imposes that a government budget must be balanced under the chosen policies. The fifth constraint captures the fact that the social planner understands that y_c will be influenced by the chosen policies through effects on parental human capital investments. After some algebra, the planner's optimal policies reduce to,

$$\begin{aligned} s* &= 1-\lambda \\ \tau* &= (1-\lambda)\tfrac{\beta}{1+\beta}. \end{aligned} \tag{10}$$

Clearly, as the degree of strategic complementarity rises the social planner finds interventions more desirable. In this simplified scenario the optimal subsidy moves one-to-one with the degree of skill substitutability in production.

While substantial interventions may occur when strategic complementarity is sufficiently intense, such changes would not affect the intergenerational earnings elasticity. This can easily be seen by writing a version of equation (8) that holds under these policies. To derive such an expression we substitute the optimal policies into the solution to problem (9), which results in:

$$w(h_c) = \left[\frac{\beta\lambda(1+\beta)}{\beta(1+\beta\lambda)}\right]^{\lambda} y_c^{1-\lambda} w(h_p)^{\lambda}. \tag{11}$$

Under this policy scenario the intercept of the intergenerational log-earnings regression would be affected, but not its slope. However, independence between policy and intergenerational elasticity is not a general result. If public investments are lump-sum rather than proportional – a perhaps more realistic representation of public schooling systems – intergenerational mobility will be affected. Augmenting our original setup so that $h_c = m + S$, we would have the following expression for a child's human capital attainment:

$$h_c = \max\left\{S, (1-\tau)\frac{\beta\lambda}{1+\beta\lambda} y_p^{1-\lambda} h_p^{\lambda}\right\}. \tag{12}$$

Among families with high-income parents the intergenerational elasticity will continue to be λ, but among low income families (those for whom $h_c = S$), the intergenerational elasticity will be zero. Clearly, lump-sum education policies would reduce the overall intergenerational earnings elasticity. Optimal policy in this setting is non-trivial, and we defer a discussion to the richer numerical analysis below.

The presence of a relationship between the IGE and public education spending can also be detected in data. For example, Blanden (2013) uses data by Barro and Lee (1994) about government education spending as a proportion of GDP (for both total and recurring expenditures) and shows that countries which devote more of their income to public spending on human capital investment tend to be more mobile.[5]

3 Evidence from the Geography of Earnings' Mobility

Geographic variation may convey valuable information about the role of technology in driving social mobility. Here we provide reduced-form evidence suggesting that regional differences in economic mobility are in fact associated to the geography of industry composition. We do so by constructing proxies of the 'average' degree of skill substitutability in production for different locations, which we then relate to local measures of inter-generational earnings

[5]Table 8 in Blanden (2013) shows the expected negative relationship between education spending and intergenerational persistence. The data allows one to take average figures from 1965 to 1969 (the primary school years for the 1960 cohort) and 1970 to 1974 (the early secondary school years) and correlate them with measures of mobility. There is no consistent pattern on the most important period of schooling, primary or secondary.

persistence. The analysis is performed at the country level.

3.1 Cross-Country Differences

The hypothesis that intergenerational persistence within a country may be related to that country's production arrangements is hard to test directly. The lack of accurate and comparable data on the lifetime incomes of parents and children for a large set of countries makes it difficult to assess the importance of different mechanisms driving cross-country variation in the intergenerational persistence of income and economic status.

To measure cross-country variation in intergenerational mobility we have gathered estimates from a large number of studies measuring the inter-generational elasticity of earnings. For some countries several measurements are available, a fact that allows us to run robustness checks of our reduced-form results.

For each country in our sample we construct a skill substitutability proxy, which captures differences in skill substitutability resulting from variation in industry composition. In practice, these proxies are constructed in two steps: (i) we devise measures of skill substitutability in production for different industries; (ii) we use STAN OECD data on sectoral value-added to weigh each industry and generate a country-specific measure of average substitutability of skills. In what follows we briefly describe how different proxies are constructed.

Industry measures of skill substitutability. Our main measure of skill complementarity utilizes information on the way workers interact during production, obtained from the O*NET database. This information consists of direct measures about the way workers skills impact output, as the O*NET reports information on occupation-specific requirements, including measures of skill substitutability in production.[6] We aggregate these measures at the industry level to create industry-specific proxies. The O*NET measures are based on workers' answers to simple questions. To be useful, the questions must capture the degree of skill-substitutability in both the manufacturing and services industries. Moreover, they must be sufficiently unambiguous in their phrasing so that they can be interpreted as direct proxies of skill substitutability. For this analysis we choose three questions that focus on the degree to which each worker's output depends on the skills of her co-workers, as well as on her own skills. We select questions that were posed directly to workers, rather than measures constructed by analysts. The first question asks workers whether they "work as a team member". The answer to this question is a simple yes or no. In contrast, the measurement scale for the remaining two questions ranges between 1 and 5. The second question is slightly more nuanced, asking workers: "How

[6] For a description of the O*NET and its precursor (the DOT) see Cortes and Gallipoli (2015).

important is independence to the performance of your current job?". The question highlights whether a "job requires developing one's own ways of doing things, guiding oneself with little or no supervision, and depending on oneself to get things done". An occupation with the latter characteristics is one in which own skills and effort are the main drivers of productivity, rather than interaction with other people; to make this proxy comparable with the other measures of substitutability, we generate a "lack of" independence variable. The third question relates to each worker's responsibility for the final outcomes and results in production. The exact question is: "How responsible are you for work outcomes and results of other workers on your current job?". This latter question measures the extent to which a particular worker's input may affect the output of co-workers.

We also assess the robustness of this measure by considering an alternative measure of skill complementarity that is based on theory.[7] Our model suggests that industries in which skill substitutability is relatively stronger are, ceteris paribus, characterized by higher wage dispersion.[8] The wage dispersion proxies are obtained using US data from the Current Population Suvery (CPS) in the year 2000. We present results using either the standard deviation or the coefficient of variation as measures of wage dispersion, but we also verify robustness using inter-percentile differences (results available from the authors). We consider both raw wages and residual wages; the latter are obtained by first regressing measures of individual wages on a set of observable characteristics.[9] The raw (or residual) wages are then grouped into industries corresponding to the ISIC classification adopted in the STAN data set. At this level of disaggregation there are 30 different industries. To obtain industry weights we take an average of each industry's value-added share over the five year interval between 2001 and 2005, as recorded in STAN data. Finally, using these weights, we build a measure of wage dispersion for each country. While this alternative measure is not used in our main analysis, it does illustrate the robustness of our reduced form results to alternative skill complementarity measures.

In summary, we use two alternative sets of skill substitutability measures (based, respectively, on O*NET measures or wage dispersion). Each set consists of various different proxies. The O*NET measures directly approximate the importance of team membership, responsibility for others' output and independence in production. As such, the main advantage of these measures is that they do not rely on economic theory, or on any assumption implicit in our model.

[7] For a detailed discussion of this approach see Bombardini, Gallipoli, and Pupato (2012; 2014).

[8] This relationship, which we also use to identify and estimate industry-specific skill complementarity, is explicitly derived in Section 4 where the structural model is described.

[9] In the first stage we control for age, education, gender, industry, State and MSA, race, veteran status and employment type.

In the subset of wage dispersion measures we include the standard deviation and the coefficient of variation for both raw and residual wages. These measures are derived from economic theory, and their use as proxies of substitutability in production is model-dependent. As such, we only use them to verify the robustness of the results based on O*NET measures.

It is important to highlight that while the proxies in the O*NET subset are negatively correlated with skill substitutability, the opposite is true for the wage dispersion proxies, which instead are increasing in skill substitutability. We first use each proxy within a subset separately; then, we perform factor analysis to identify the principal component driving the measures within each subset. This gives us a unique proxy of skill substitutability within each industry, which should in principle be more precise, as it is generated by aggregating the noisy information of all proxies in a given subset.[10]

Country-level measures of the inter-generational elasticity of earnings. Estimates of the IGE are available for several countries and periods. However, the methods and data used to obtain such estimates vary across studies. To account for this problem, we construct different samples of countries for which we observe IGE. The first sample only includes IGE estimates for nine countries corresponding to the preferred sample listed in Table 1 of Corak (2006) — namely Canada, Denmark, Finland, France, Germany, Norway, Sweden, UK and US. We call this sample the 'core' sample. These countries are chosen because a large number of reliable, and comparable, estimates of the IGE exists for them. Multiple measurements allow researchers to form a fairly good idea about the true IGE value in these countries. An additional advantage of this small set of nine countries is that Corak's study provides a set of 'low end' and 'high end' estimates of the IGE for these countries. This additional information can be used to verify the robustness of the cross-country results.

To gauge the sensitivity of our findings we also extend the core sample using observations for five more countries, namely Australia, Japan, Korea, Netherlands and Switzerland. These countries are not part of the preferred subset in Corak (2006) because fewer estimates of the IGE are available for them. Nonetheless the IGE estimates are obtained using data sets and methods which are fairly comparable to those used in the core sample. Adding these countries increases the sample size, but it also adds noise as IGE estimates are likely to be less accurate. All the alternative IGE samples are reported in Table (1).

Findings of cross-country analysis. In what follows we overview results from least-square

[10]A restriction, satisfied by all the measures we use, is that different proxies within a subset must positively correlate across different occupations and industries, indicating that they similarly co-vary with skill substitutability. This restriction makes interpretation of the common components relatively straightforward.

Table 1: Inter-generational elasticity of earnings (IGE) in different countries: different samples.

	core	core (low end)	core (high end)	core + 5
	(1)	(2)	(3)	(4)
Country				
Australia				0.26
Canada	0.19	0.16	0.21	0.19
Denmark	0.15	0.13	0.16	0.15
Finland	0.18	0.16	0.21	0.18
France	0.41	0.35	0.45	0.41
Germany	0.32	0.27	0.35	0.32
Japan				0.34
Korea				0.25
Netherlands				0.23
Norway	0.17	0.15	0.19	0.17
Sweden	0.27	0.23	0.3	0.27
Switzerland				0.46
UK	0.5	0.43	0.55	0.5
USA	0.47	0.4	0.52	0.47

regressions of each country's IGE on proxies of skill-substitutability. We begin by focusing on the small (core) sample of nine countries for which we have fairly accurate and consistent measurements of the IGE. The first three columns in Table (2) report the estimated changes in the level of IGE associated to a one-standard-deviation increase in the skill substitutability proxy.

Each column refers to a regression estimated using a different set of IGE measures: in the first column (core) we use the preferred estimates of Corak (2006) for the nine countries in the sample; the next two columns report results of regressions in which we use either the lower end or the higher end estimates for the same core sample. The last column reports results from an expanded sample in which we add five extra countries to the analysis. The top panel (A) reports results based on O*NET measures of complementarity, while the bottom panel (B) performs the same analysis using complementarity measures based on industry-specific wage dispersion. Given the fact that O*NET proxies are increasing in complementarity while dispersion proxies are decreasing, the sign of the estimated coefficients are not the same in Panel (A) and Panel (B), however the magnitudes are comparable to each other because they are based on standardized proxies.

The results are quite striking, especially if one considers the small sample sizes. First, in all the samples we find significant variation in the conditional mean of the IGE as the cross-

Table 2: Estimated change in IGE associated to a one-standard-deviation decrease in the skill-substitutability proxy. Top panel (A): proxies based on ONET measures. Bottom panel (B): proxies based on wage dispersion. Columns correspond to different samples ('core' sample in first column; extended sample in second column). Standard errors in parenthesis.

Panel (A)	IGE measures			
	core	core (low end)	core (high end)	core + 5
skill substitutability proxy	(1)	(2)	(3)	(4)
ONET proxies (decreasing in substitutability)				
team member	-.068	-.056	-.073	-.055
	(.032)	(.027)	(.036)	(.025)
responsibility	-.090	-.076	-.097	-.041
	(.037)	(.031)	(.041)	(.031)
independence	-.081	-.068	-.087	-.054
	(.027)	(.022)	(.030)	(.028)
common component[b]	-.086	-.072	-.092	-.054
	(.033)	(.028)	(.037)	(.030)

Panel (B)	IGE measures			
	core	core (low end)	core (high end)	core + 5
skill substitutability proxy	(1)	(2)	(3)	(4)
Wage dispersion proxies[a] (increasing in substitutability)				
S.D. of raw wages	.108	.091	.117	.090
	(.029)	(.025)	(.032)	(.019)
C.V. of raw wages	.100	.084	.109	.080
	(.033)	(.029)	(.036)	(.021)
S.D. of residual wages	.107	.090	.116	.083
	(.028)	(.024)	(.030)	(.017)
C.V. of residual wages	.095	.079	.102	.068
	(.031)	(.026)	(.034)	(.017)
common component[b]	.104	.087	.113	.082
	(.031)	(.027)	(.034)	(.019)

Notes
a Common component estimated using skill substitutability proxies listed in each panel.
b C.V.=coefficient of variation; S.D.=standard deviation.

country substitutability changes. Second, the estimated conditional mean differences are sizeable: a one-standard-deviation change in the skill-substitutability proxy induces a difference in IGE of between 5 and 10 basis points, depending on the sample and the measure of skill substitutability. These changes are not small, considering that most IGEs have a size somewhere between 15 and 50 basis points (see Table 1)[11]. Third, the conditional effects on the IGE are fairly similar and do not depend on (i) the specific IGE sample we use (core, lower or higher estimates) or, (ii) the specific skill-substitutability proxy.

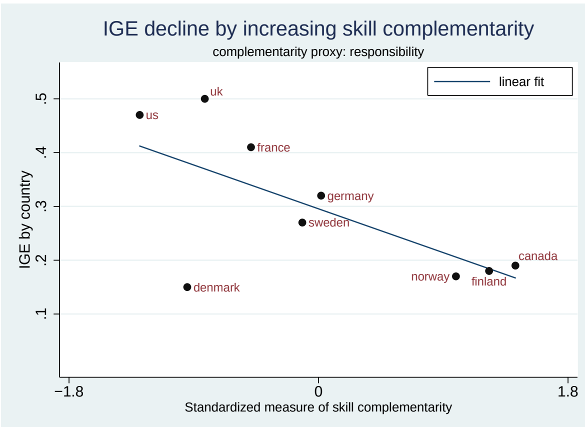
To make the latter point more transparent we report in Figures 2 and 3 the scatter plots of IGEs, for the 'core' and extended country samples, versus the skill substitutability indices. In both figures we superimpose a linear fit line. The plots in Figure (2) and (3) are based on different sets of skill substitutability proxies. As we change the country samples — or the way substitutability is measured — similar patterns are detected. Observations consistently line up around the fit lines and, crucially, countries with relatively stronger skill substitutability in production exhibit a higher IGE. One notable exception is Denmark, which appears to have a much lower IGE than would be predicted by its industry composition.

An interesting feature of the plots in Figure (3) is that they offer an alternative view of the so-called 'Great Gatsby' relationship between intergenerational income persistence and inequality. Unlike the usual accounts of this relationship, the country-specific proxies of wage inequality in Figure (3) are obtained from weighted averages of US industry wage dispersion measures, where weights change by country and correspond to industry shares. The fact that a clear relationship is apparent, despite the use of wage dispersion measures from a single country, suggests that industry composition may be key for that relationship. This evidence also hints at the possibility that patterns of industry wage dispersion may be fairly stable across countries, a fact that would be consistent with the hypothesis that industry-specific skill complementarity does not vary dramatically between countries. We explore this possibility in more detail in Appendix A.3, where we show how the ranking of both earnings and wage dispersion across industries is remarkably stable across countries and over the time period considered.

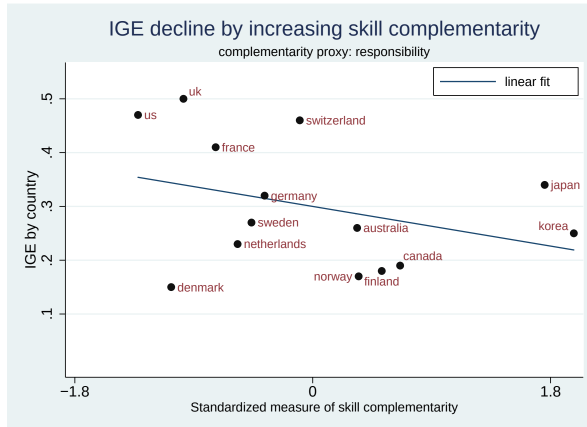
When we extend the core sample by including Australia, Japan, Korea, Netherlands and Switzerland we end up with a set of 14 IGE measures corresponding to those listed in column (4) of Table (1). The last column of Table (2) reports estimated slope coefficients of the univariate regression using the extended sample of IGEs and different measures of country-specific skill substitutability. Significance levels are consistent with 'core' sample results, and magnitudes are comparable, albeit slightly lower. The scatter plots, in the right panels of Figures (2) and (3), confirm a robust correlation between measures of skill substitutability and IGE levels.

[11] The R^2 values of these univariate regressions are relatively large, often exceeding 40%.

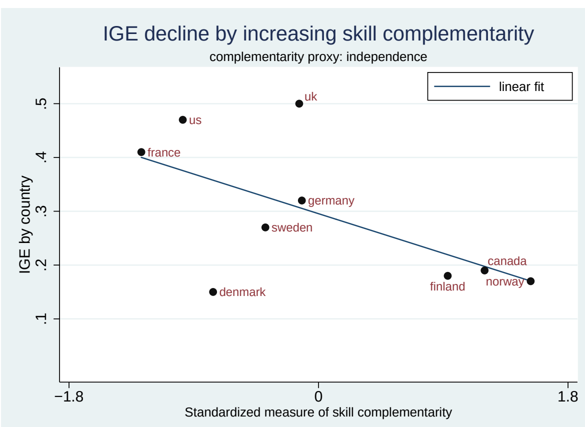
Figure 2: IGE vs average skill substitutability, in both the core IGE sample (left column) and the extended sample (right colum). ONET proxies: (top row) responsibility; (middle row) independence; (bottom row) common factor estimated from three separate ONET proxies (responsibility, independence, team member).



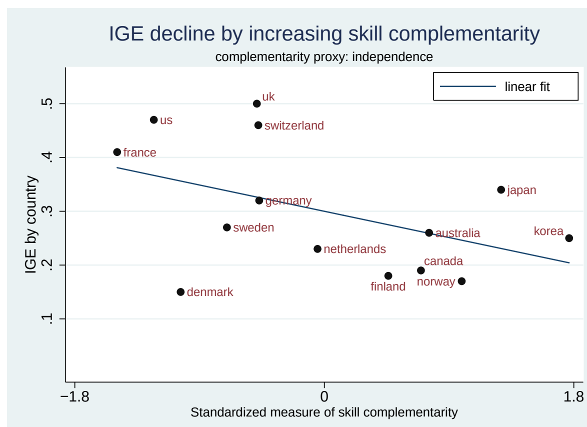
(a) core IGE sample, responsibility proxy



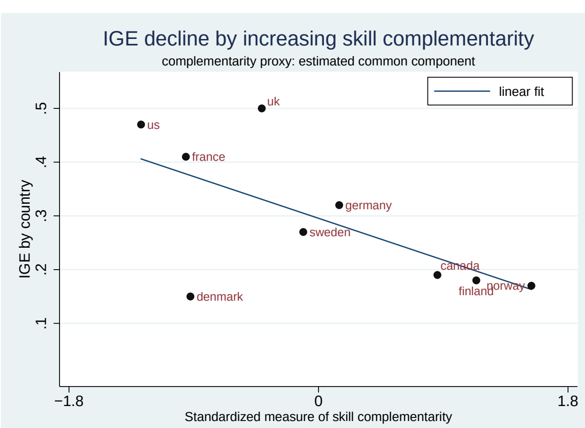
(b) extended IGE sample, responsibility proxy



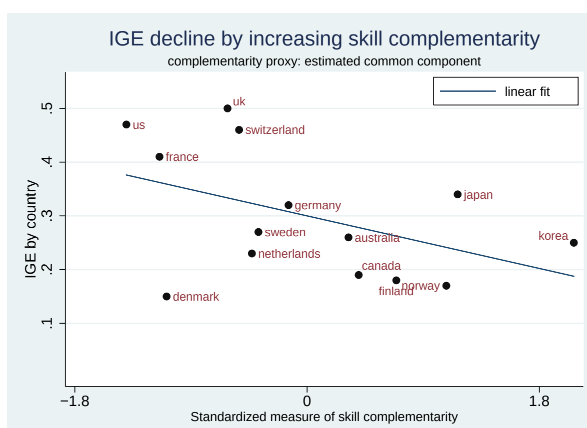
(c) core IGE sample, team member proxy



(d) extended IGE sample, team member proxy

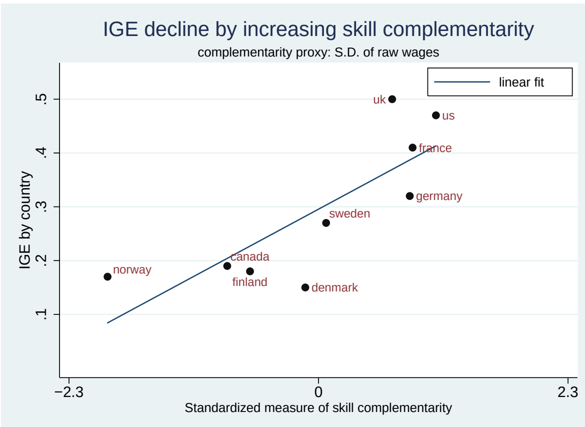


(e) core IGE sample, common component proxy

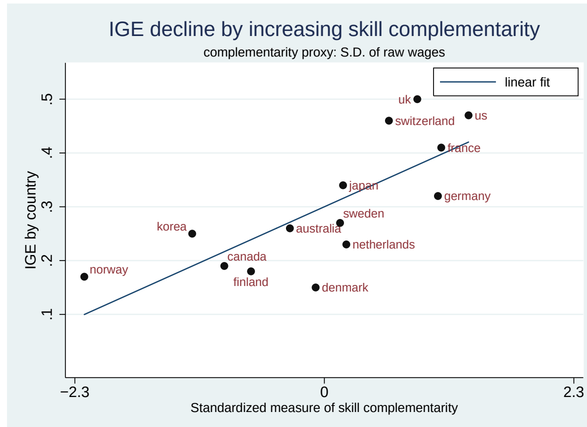


(f) extended IGE sample, common component proxy

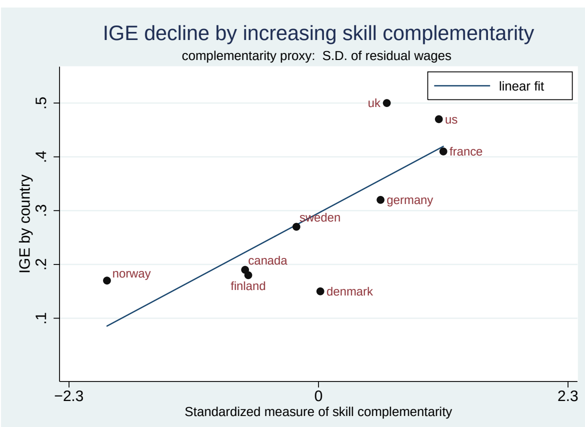
Figure 3: IGE vs average skill substitutability, in both the core IGE sample (left column) and the extended sample (right colum). Wage dispersion proxies: (top row) S.D. of raw wages; (middle row) S.D. of residual wages; (bottom row) common factor estimated from four separate wage dispersion proxies.



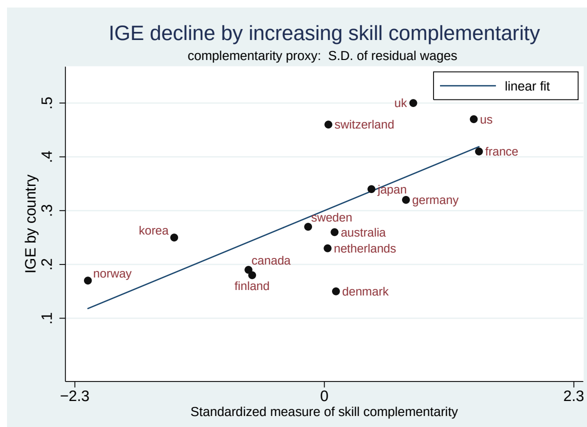
(a) core IGE sample, S.D. of raw wages proxy



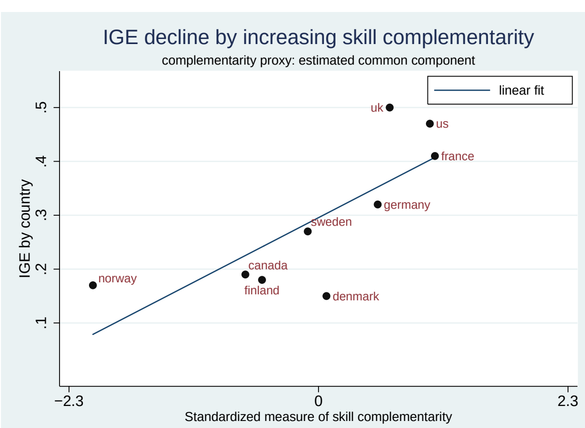
(b) extended IGE sample, S.D. of raw wages proxy



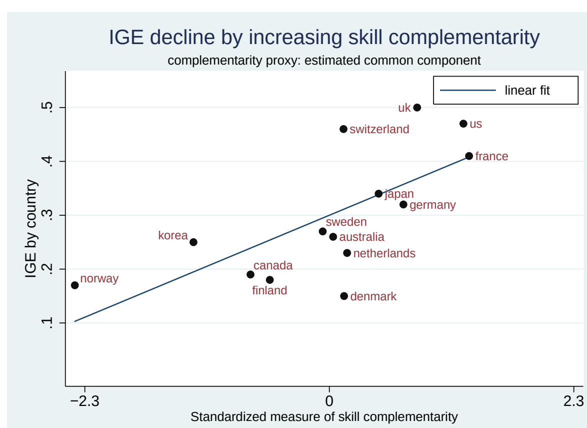
(c) core IGE sample, S.D. of residual wages proxy



(d) extended IGE sample, S.D. of residual wages proxy



(e) core IGE sample, common component proxy



(f) extended IGE sample, common component proxy

4 Structural Analysis

We study a model that has two key features. Firstly, the aggregate production function in the model exhibits the same type of skill complementarity as our simple model did, namely a form of skill complementarity that induces strategic complementarity in parental human capital investments. The production function will thus need to have a form somewhat different than the CES production functions usually utilized in macroeconomics. In our production function an individual worker's marginal product increases when the average skill level among her co-workers rises. In contrast, in some common CES production functions, e.g. Goldin and Katz (1995), adding low skilled workers (hence decreasing the average skill of co-workers) increases other workers' productivity.

Secondly, we require a model that we can map to data in a logical way. For this purpose we posit a technology with multiple production sectors, which correspond to the 2-digit industry aggregates in OECD STAN data. Each industry has its own aggregate input of human capital, and its own unique degree of skill complementarity, which we explicitly estimate. Adopting this approach allows us to design simple counterfactual scenarios and quantify the importance of skill substitutability in a clear and straightforward way. The share of GDP generated by each industry is determined by a single parameter: therefore, one can simply vary the sector share parameters to reflect the industry composition– and thus average skill complementarity –of a different country.

We study a steady-state overlapping generations economy. Each generation consists of a unit continuum of households, who live for two periods. During the first period of life households are non-autonomous, their only activity being the acquisition of human capital. During the second period of life households earn wages, which depend on human capital levels, and decide how much to invest in the human capital of their children.

The model explicitly accounts for the two-way relationship between distribution of skills and industrial composition. The aggregate skill distribution, and the allocation of skills across industries, are shaped by parental decisions. These distributions determine the incomes of the current generation of workers, and hence the human capital investments they make for their children. In the long-run the economy reaches an equilibrium, in which both cross-industry allocations and the skill distribution are stationary. Differences in primitive structural parameters across countries, which we estimate, determine differences in these steady state outcomes and, as a consequence, differences in intergenerational mobility.

4.1 Model

Production

The production side of the economy consists of N sectors, the outputs of which are aggregated into a final consumption good. The inputs to the production of any intermediate good y_n are the industry-specific capital k_n, and industry specific human capital ℓ_n:

$$y_n = k_n^{\alpha_n} \ell_n^{1-\alpha_n}. \tag{13}$$

The capital share is allowed to vary by industry, as observed in data. This means that differences in capital intensity may affect worker productivity across industries. The industry-specific human capital input depends on four factors: the number of workers in the industry, the workers' individual skill levels, the workers' idiosyncratic productivity realizations, and the industry-specific substitutability of skills. Crucially, we allow for strategic complementarity in human capital investments. The specification of ℓ_n is

$$\ell_n = \left(\int_{I_n} z(i) h(i)^{\lambda_n} di \right)^{\frac{1}{\lambda_n}}, \tag{14}$$

where I_n is the set of workers in industry n, $z(i)$ is the realized productivity of worker i, $h(i)$ is the skill level of worker i, and λ_n reflects skill substitutability in this industry. As the mass within the set I_n grows, so does the human capital input of industry n; adding an additional high-skilled worker increases ℓ_n by more than an additional low-skilled worker, and more so if skills are easily substitutable.

The timing is such that workers' idiosyncratic productivity shocks are observed after they have been allocated to an industry. This results in the distribution of realized productivity shocks in every industry being the same as the aggregate distribution of productivity shocks, which is known. Therefore, relaxing this assumption would not alter the nature of the equilibrium sorting of workers to industries[12]. One interpretation of the productivity shocks is a lifetime labor supply shock, which does not affect the skill attainment, but rather the lifetime labor supply.

[12] Relaxing this assumption would introduce a non-uniqueness problem in the sorting of workers across firms. However, this non-uniqueness would affect only the distribution of $z(i)$ across industries, rather than the distribution of the $h(i)$'s. For example, if it were optimal for an industry to hire 10 units of type $\hat{h}$ workers, this would be the case even if the timing of productivity realizations was relaxed. The non-uniqueness would arise because 10 units of $\hat{h}$ could be attained either by hiring 10 workers for whom $z(i) = 1$, or five worker for whom $z(i) = 2$, or any other such combination.

Output of the final consumption good is a Cobb-Douglas aggregate of all intermediate inputs:

$$Y = \prod_{n=1}^{N} y_n^{\gamma_n}. \tag{15}$$

The weights γ_n reflect the relative size of each industry, where $\sum_{n=1}^{N} \gamma_n = 1$. These weights can be directly measured from data on output shares. Adjustment of these weights to reflect observed international differences is one of the key sources of variation in the counterfactual experiments.[13] We interpret the weights γ_n as arising from a combination of resource endowments, historical occurrences, geography, climate, and similar long-term characteristics, which together affect relative advantage, industrial composition and aggregate technology.[14]

The labor input within each industry has the same CES specification as the aggregate production function used in the two-period analytical example. However, unlike the simple model, the set of workers within a given industry is endogenous because workers can sort freely across industries. This results in a worker-industry matching problem, similar in nature to the setting studied by Pycia (2012). In such a setting it can be difficult to characterize a stable match equilibrium because each worker's productivity depends on that of co-workers, and differently so across industries. To make this problem tractable we make two simple assumptions that allow us to write the equilibrium allocation as the solution of a standard Kuhn-Tucker program. The two assumptions are (1) that skill levels change in discrete steps, and (2) that skill levels are finite in number. As shown below, this allows us to transform the problem into a simple constrained maximization problem.

Let the finite set of possible human capital attainments be H, and let the measure of workers with attainment h in industry n be $q_n(h)$. Then the industry production function 13 can be rewritten as

$$\ell_n = \left(\int_H \int_Z z dF(z) q_n(h) h^{\lambda_n} dh \right)^{\frac{1}{\lambda_n}}, \tag{16}$$

where we integrate over the set of skill levels and productivity realizations, rather than over the set of workers within the industry. Note that our assumption that productivity risk is realized after matching has allowed us to simply weight the measure of workers by their average productivity. Viewing the aggregate production technology as operated by a competitive rep-

[13]This specification of aggregate output, while restricting the elasticity of substitution, has the advantage of making the mapping of the model to data extremely transparent.

[14]Over time specialization patterns may also respond to changes in the characteristics of the working population. We discuss the co-determination of industry structure and human capital after reviewing our main results.

resentative firm, the appropriate profit maximization problem is:

$$\begin{aligned} \max_{\{q_n(h)\}} \quad & Y(\{q_n(h)\},\{k_n\}) - \textstyle\sum_{n=1}^{N} \int_H \int_Z z w(h,n) q_n(h) dF(z) dh - \sum_{n=1}^{N} (r+\delta) k_n \\ & s.t. \\ & \{q_n(h)\} \geq \mathbf{0}, \\ & \{k_n\} \geq \mathbf{0}, \end{aligned} \tag{17}$$

where $w(h,n)$ is the wage per unit of human capital of skill-level h in industry n. The vector $\{q_n(h)\}$ contains measures of human capital inputs for each industry-skill pair, and the vector $\{k_n\}$ contains all industry specific physical capital inputs.

First-order optimality conditions for profit maximization include complementary slackness conditions for human capital measures $q_n(h)$:

$$q_n(h)\left[\frac{\partial Y(\{q_n(h)\},\{k_n\})}{\partial q_n(h)} - w(h,n)\right] = 0. \tag{18}$$

These optimality restrictions state that workers with skill level h are either paid their marginal product within an industry, or there is a measure zero of them working in that industry. The wage paid per unit of human capital of skill level h in an industry n is:

$$w(h,n) = Y\gamma_n(1-\alpha_n)\frac{1}{\lambda_n}\left\{\frac{h^{\lambda_n}}{(\int_H q_n(h)h^{\lambda_n}dh)}\right\}, \tag{19}$$

where the right-hand-side is the marginal product from equation (18). This labor demand holds whenever $q_n(h) > 0$, and implies that the variance of wages within industry n is an increasing function of λ_n. In fact, one can show that $Var_n(\ln w(h)) = \lambda_n^2 \cdot Var(\ln h)$. Firms also hire capital services optimally, thus sectoral inputs satisfy:

$$\frac{\partial Y(\{q_n(h)\},\{k_n\})}{\partial k_n} - (1-\tau_k)r - \delta = 0, \tag{20}$$

where δ is the capital depreciation rate. The real interest rate r is assumed to be exogenously determined in world capital markets, thus equilibrium capital inputs are easily computable from these equations.

Households

The household side of the economy is similar to that in Restuccia and Urrutia (2004). A household's adult wage depends on two state variables: the endowment of skills, h, and the realized idiosyncratic shock, z. The idiosyncratic shock is distributed log-normally with mean μ_z and variance σ_z^2,

$$\ln z \sim N(\mu_z, \sigma_z^2). \tag{21}$$

The variance σ_z^2 is set to replicate the degree of idiosyncratic risk observed in U.S. earnings data, as described in Section (4.6). The mean is set so as to normalize average productivity to unity, i.e. $\mathbb{E}[z] = 1$.

A child's achievement, $h' = g(\theta', m + s)$, depends on her endowment of heritable traits $\theta' \in \Theta$, parental investments m and public investments s. We use a one-period stand-in to approximate the complicated dynamic skill-formation technology described in (e.g. Cunha, Heckman, and Schennach, 2010; Del Boca, Flinn, and Wiswall, 2014; Abbott, 2015). We assume the same technology as that employed in Restuccia and Urrutia (2004) and earlier by Becker (1981):

$$h' = \theta'(m + s)^{\psi}. \tag{22}$$

Investments are units of resources. Heritable traits are persistent across generations and follow a mean-zero AR(1) process, as in Solon (2004):

$$\begin{aligned} \ln(\theta') &= \rho \ln(\theta) + \eta \\ \eta &\sim N(0, \sigma_\eta^2). \end{aligned} \tag{23}$$

All exogenous intergenerational persistence is driven by this component, as idiosyncratic income risk (z) is iid across generations.

Utility from consumption is of a CRRA form, and the altruism weight a parent puts on their child's wellbeing is denoted as β. To capture the progressiveness of U.S. tax policies, we implement a proportional wage tax, τ, and transfer a proportion of revenue back in lump-sum fashion, T. Parents may transfer wealth $a' \geq 0$ to their children, in addition to investing m in their human capital. Given this structure, the parental decision problem can be represented recursively as:

$$
\begin{aligned}
V(a,h,\theta',z) &= \max_{c,m}\left\{\tfrac{c^{1-\sigma}}{1-\sigma}+\beta\mathbb{E}\left[V(a',h',\theta'',z')|\theta'\right]\right\} \\
s.t. & \\
c+m+a' &= zW(h)(1-\tau)+T+a(1+r) \\
h' &= g(\theta',m+s) \\
\ln(\theta'') &\sim N(\rho\ln(\theta'),\sigma_\eta^2) \\
\ln(z') &\sim N(\mu_z,\sigma_z^2). \\
a' &\geq 0 \\
W(h) &= \max_n\{w(h,n)\}.
\end{aligned}
\tag{24}
$$

Each parent has full information about his child's inherited traits, but does not know what the realized productivity shock z' will be nor the inherited traits of the grandchild θ''. The final equation of the problem is an implicit labor supply condition, which says that a worker with skill h will always work in the industry that rewards her skill the most.

Government

Any tax revenues in excess of the lump-sum transfers are spent on non-valued expenditure, G. The government budget constraint is,

$$
G=\tau\int w(h(i))z(i)di+\tau_k\sum_{n=1}^{N}rk_n-T-s. \tag{25}
$$

4.2 Equilibrium

We use the notion of stationary competitive equilibrium[15] and define it as a collection of:

(i) decision rules $\{c(a,h,\theta',z),m(a,h,\theta',z),a'(a,h,\theta',z)\}$ for consumption, human capital investments and asset transfers, and the value function $V(a,h,\theta',z)$;

(ii) Aggregate industry specific human capital attainment measures $\{q_n(h)\}$;

(iii) Wages $\{w(h,n)\}$;

(iv) and state-space measure μ; such that

[15] This definition is simplified because of the open-economy assumption with outside determination of the real interest rate.

1. The decision rules solve the household optimization problem (9), and $V(a, h, \theta', z)$ is the associated value function.

2. The representative firm optimally hires human and physical capital, thus equations (18) and (20) hold.

3. Each skill and industry specific labor market clears

$$\begin{aligned} \textstyle\sum_{n=1}^{N} q_n(h) &= \int\limits_{A\times H\times\Theta\times Z} 1_h d\mu \ \forall \ h \in H, \\ \text{and } 0 &= q_n(h)\left[w(h,n) - \partial Y/\partial q_n(h)\right] \end{aligned} \tag{26}$$

where 1_h is an indicator function for the state variable h.

4. The government budget constraint in equation (25) holds.

5. Individual and aggregate behaviors are consistent: the measure μ is the fixed point of $\mu(S) = Q(S, \mu)$ where (i) $Q(S, \cdot)$ is a transition function generated by the individual decision rules and the exogenous laws of motion for θ' and z; and (ii) S is the generic subset of the Borel-sigma algebra $\mathcal{B}_S$ defined over the state space $A \times H \times \Theta \times Z$.

4.3 Worker-Industry Matching in Equilibrium

The worker-industry sorting that occurs in equilibrium hinges on skill substitutability differences. Matching is positively assortative between a worker's human capital level and the industry-specific degree of skill substitutability. That is, the most skilled workers sort into industries where skills are the most substitutable because it is in these industries that they capture the largest returns. The choices of high skill workers push lower skill workers towards industries with more skill complementarity. We formalize these matching patterns in a proposition stating that, for any two workers with diverse skill levels employed in different industries, the higher skilled worker must be employed in the industry for which λ is larger.

Proposition 1 *Suppose workers i and j have skill levels h_i and h_j, where $h_i > h_j$. If worker i is in industry 1 and worker j is in industry 2, then $\lambda_1 \geq \lambda_2$.*

Proof. In equilibrium workers choose industries where their wage will be the highest. Then $w(i, 1) \geq w(i, 2)$, and $w(j, 1) \leq w(j, 2)$. Therefore,

$$\frac{w(i,1)}{w(j,1)} \geq \frac{w(i,2)}{w(j,2)}.$$

Using the wage equations (labor demand) this implies

$$\left(\frac{h_i}{h_j}\right)^{\lambda_1} \geq \left(\frac{h_i}{h_j}\right)^{\lambda_2}.$$

■

The proof shows how the ratio of the marginal products of high to low skilled workers will always be larger in industries where skills are more substitutable. Because this is true, in any counterfactual case where the low skilled worker is in the high substitutability industry, greater efficiency could be attained by switching the two workers.

4.4 Benchmark Model: Parameterization and Properties

In what follows we describe how the numerical counterpart of the model is parameterized. We do so by focusing, in turn, on the production side, the household side and on tax and benefits parameters. Tables (3) and (4) report parameter values. We also report a summary of the moments targeted in our calibration in Table (3). In Section (4.8) we describe some key properties of the benchmark model.

4.5 Production parameters

Industry-specific physical capital. The quantity of physical capital in each industry depends on the capital share α_n in that industry, and on the (exogenous) gross return on capital $r+\delta$. We set the real interest and depreciation rates so that their annualized values are 3.5% and 6.0%, respectively. Each industry-specific share of output paid to capital is measured using OECD STAN data and is set equal to α_n.

Intermediate goods aggregation. The relative value of output derived from each industry is equal to its weight γ_n. Thus, the aggregation weights can be parameterized by setting them equal to the share of total output (value added) attributed to each industry. To this purpose we compute, for each industry, the average of the value-added shares observed in STAN OECD data over the five years from 2001 to 2005.

Industry-specific elasticity of substitution. We use our O*Net indicators in the parametrization of complementarity. More specifically, we adopt an indirect inference approach and use the O*Net measure as a noisy indicator of variation in λ_n by positing the relationship,

$$\lambda_n = a_0 + a_1 \cdot \text{O*Net}_n. \tag{27}$$

The identification strategy relies on model implications for the relationship between a_1, a_2 and earnings variation. In particular, we use information about: (i) the degree of total cross-sectional variation in earnings; (ii) the relationship between intra-industry earnings variation and our O*Net indicators. To understand how variation in these moments identifies a_1 and a_2 consider the following expression for the variance of log earnings within industry n:

$$Var_n(\ln y) = Var(\ln z) + (a_0 + a_1 \cdot \text{O*Net}_n)^2 \cdot Var_n(\ln h) \tag{28}$$

This variance decomposition is derived from the labor demand (wage) equation (19) with λ_n replaced by the linear specification (27).[16]

The parameters (a_0, a_1) are jointly identified by overall earnings dispersion and by differences in within-industry dispersion. The variance of earnings in every industry rises if a_0 increases. Thus a_0 can be identified by targeting overall earnings dispersion produced by the model to what is observed from data. Crucially, the log earnings variation to be matched is that of lifetime earnings, which Bowlus and Robin (2012) suggest is 30% lower than its cross-sectional counterpart. Estimates of the cross-sectional variance of log earnings are generally close to 0.6 (e.g. Heathcote, Storesletten, and Violante, 2010), thus we target a standard deviation of log lifetime earnings equal to 0.42.

To identify a_1 we use information about the relationship between within-industry wage dispersion and industry O*Net scores, as described by equation (28).[17] From CPS data we compute the standard deviation of log earnings within each industry, and then regress these standard deviations on the industry-level O*Net scores. We find that the slope coefficient from this regression is equal to 0.56, which we use as a target. Because O*Net scores are a principle component factor with unit variance, the regression coefficient is also the covariance between O*Net scores and within industry earnings standard deviations. We denote this $\text{cov}(\sigma_n(\ln y), \text{ONet}_n)$ in Table (3).

In the spirit of indirect inference, we use a nested fixed-point algorithm to pin down a_0 and a_1 (as well as the other calibrated parameters). That is, we guess values for a_0 and a_1 and simulate the model. If, for example, the relationship between O*Net and within industry wage

[16]Write down the logarithm of the wage equation as

$$\ln(z \cdot w(h)) = \ln\left[\gamma_n \frac{1-\alpha_n}{\lambda_n} Y \left(\frac{1}{\ell_n}\right)^{\lambda_n}\right] + \ln z + \lambda_n \cdot \ln h.$$

Then the variance of wages within industry n is $Var_n(\ln y) = Var(\ln z) + \lambda_n^2 \cdot Var(\ln h)$.

[17]Note that also a_0 has an effect on within-industry dispersion. Hence, there may in principle be different combinations of (a_0, a_1) which fit the data equally well. For our purposes any such combination is sufficient to recover the parameter of interest λ_n.

dispersion is too weak (or strong) in the simulated data we would increase (or decrease) a_1 and re-simulate the model.

4.6 Household parameters

Preferences. We set the risk aversion parameter to $\sigma = 2$, and the discount factor to $\beta = 0.5$. The value of the discount factor reflects the time gap between outcomes of children and parents. Based on a 25 year gap, the annualized discount factor implied by our parametrization is just above 0.97.

Idiosyncratic income risk. Storesletten, Telmer, and Yaron (2004) suggest that post market-entry factors account for about 40% of income variation in U.S. data, which we adopt as a target. Because income risk is orthogonal to other sources of income variation, this target can be achieved by choosing the variance of income shocks σ_z^2 appropriately. Finally, given σ_z^2, the mean of log income risk, μ_z, can be set so that the mean of the *level* of z is unity. In other words, we choose σ_z^2 such that the ratio $\sigma(\ln(z))/\sigma(\ln(y)) = 0.4$, and choose μ_z such that $E[z] = 1$.

Human capital production. The skill formation technology is specified as in Restuccia and Urrutia (2004):

$$h' = \theta'(m + s)^{\psi}. \tag{29}$$

The elasticity of human capital with respect to expenditures, determined by ψ, regulates how much parents are willing to spend on their child's human capital. The data moment we employ to identify ψ is the proportion of GDP spent on education by private households. According to OECD data this fraction was 2.3 in 2010.

Transmission of heritable traits. The degree of persistence in heritable traits, ρ, influences the equilibrium level of intergenerational income mobility. If heritable traits are highly persistent, then the relative magnitude of returns to human capital investments is similar for parents and their children, as described in equation (29).

In order to identify the persistence of heritable traits our model targets the intergenerational persistence of earnings in the US. Since we work in a stationary environment we simply match the intergenerational correlation of earnings, which Jantti, Bratsberg, Røed, Raaum, Naylor, Osterbacka, Bjorklund, and Eriksson (2006) estimate to be 0.357.[18]

To parameterize the variance of the heritable trait shock, σ_η^2, we employ information from

[18] In our quantitative experiments we scale up intergenerational correlations by 1.32 to convert them to elasticities. This accounts for changes in the variance of earnings over time.

income quintile transition matrices. Jantti, Bratsberg, Røed, Raaum, Naylor, Osterbacka, Bjorklund, and Eriksson (2006) consider a measure of mobility based on the trace of a $(k \times k)$ transition matrix, P_k:

$$M_T = \frac{k - tr(P_k)}{k - 1}. \tag{30}$$

This measure, estimated to be 0.86 for U.S. males and 0.93 for U.S. females, gets larger as it becomes more likely for a child to enter a different income quintile than their parent. Substantially different IGEs can be generated while holding the diagonal of the transition matrix constant by adjusting the dispersion of the off-diagonal elements. Thus, the statistic M_T provides identifying information that is distinct from the IGE. If the persistent ability levels in our model are relatively dispersed then it is much less likely that an idiosyncratic shock will transit a child to a different income quintile than their parent. However, if ability levels are relatively similar, then idiosyncratic shocks can easily generate quintile transitions. Thus, given ρ and σ_z^2, the variability in heritable traits can be identified by matching $M_T = 0.89$.

4.7 Government parameters

We set the marginal tax rate to $\tau = 0.296$, which is the percentage of U.S. labor costs paid as either income tax, payroll tax, or social security contributions, as reported by the OECD (see Table 7). Like in Abbott, Gallipoli, Meghir, and Violante (2013) we calibrate the lump-sum tax rebate so as to match the progressiveness of U.S. tax policy:

$$\frac{Var\left[\ln\left((1-\tau)zW(h) + T\right)\right]}{Var\left[\ln(zW(h))\right]} = 0.61. \tag{31}$$

We calibrate s to match the 5.5% share of GDP spent publicly on education in the US.[19] Finally, the tax rate on capital income is set to $\tau_k = 0.4$ (see McDaniel, 2014).

4.8 Properties of the Benchmark Model

Table (3) presents a summary of the parameters other than industry level parameters. Table (4) summarizes the industry level parameters γ_n, α_n and λ_n, by industry. The lowest estimated λ is 0.228, and the highest is 0.970. Among the low substitutability industries are various types of manufacturing and agriculture[20], high substitutability sectors include education, health care and

[19] Table (7) in the Appendix provides details on education spending by country as reported by the OECD.

[20] Employment in US agriculture is often in large scale farms, which are organized like industrial establishments.

Table 3: Parameter values: non industry-specific parameters.

	Parameter	Value
Fixed parameters		
Intergenerational Discount Factor	β	0.5
CRRA Parameter	σ	2.0
Net Annualized Interest Rate	r	0.035
Annualized Depreciation Rate	δ	0.06
Calibrated parameters		
Idiosyncratic Risk Variance	σ_z^2	0.070
Idiosyncratic Risk Mean	μ_z	-0.035
Heritable Trait Persistence	ρ	0.429
Heritable Trait Variation	σ_η^2	0.362
Human Capital Production Weight	ψ	0.254
Lump Sum Transfer	T	471.5
Substitution Parameter Constant	a_1	0.504
Substitution Parameter Slope	a_2	1.801

Targeted Moments	Target Value	Model Value
$\sigma(\ln y)$	0.42	0.42
cov $(\sigma_n(\ln y), \text{ONet}_n)$	0.56	0.57
$\sigma(\ln z)/\sigma(\ln y)$	0.4	0.4
$E[z]$	1.0	1.0
% of GDP as Private Edu Spending	2.3	2.3
Inter-Generational Earnings Elasticity	0.47	0.47
M_T (Earnings Trans. Matrix Trace)	0.89	0.90
Ratio of Post-Tax to Pre-Tax Earnings	0.61	0.61

Table 4: Parameter values: industry-specific production parameters.

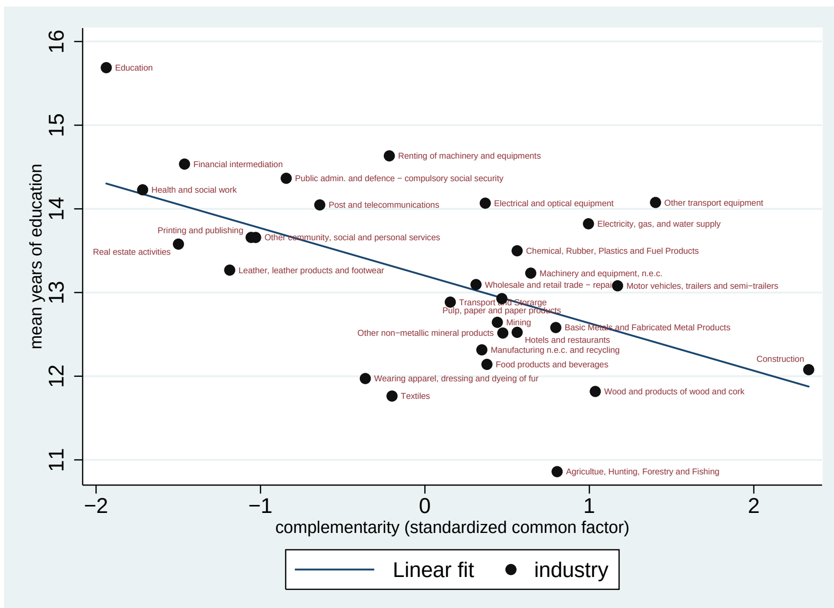
Industry	γ_n Industry Share	α_n Capital Share	λ_n Estimated Complementarity
Agriculture, Hunting, Forestry and Fishing	0.0102	0.6798	0.235
Basic Metals and Fabricated Metal Products	0.0137	0.276	0.339
Chemical, Rubber, Plastics and Fuel Products	0.0283	0.5506	0.524
Construction	0.0471	0.3364	0.491
Education	0.051	0.0847	0.970
Electrical and optical equipment	0.0167	0.0443	0.521
Electricity, gas, and water supply	0.0171	0.7094	0.458
Financial intermediation	0.0807	0.4726	0.842
Food products and beverages	0.0154	0.5724	0.270
Health and social work	0.0653	0.1862	0.955
Hotels and restaurants	0.0289	0.3825	0.273
Leather, leather products and footwear	0.0002	0.1633	0.538
Machinery and equipment, n.e.c.	0.0089	0.2689	0.435
Manufacturing n.e.c. and recycling	0.0085	0.3445	0.355
Mining	0.0125	0.6802	0.446
Motor vehicles, trailers and semi-trailers	0.0104	0.3107	0.228
Other community, social and personal services	0.0415	0.3894	0.670
Other non-metallic mineral products	0.0037	0.3573	0.506
Other transport equipment	0.0061	0.2907	0.350
Post and telecommunications	0.0305	0.4952	0.316
Printing and publishing	0.0146	0.2183	0.756
Public admin. and defence - social security	0.0792	0.204	0.694
Pulp, paper and paper products	0.0047	0.4002	0.390
Real estate activities	0.1132	0.9514	0.920
Renting of machinery and equipments	0.1294	0.3247	0.656
Textiles	0.0023	0.2079	0.397
Transport and Storarge	0.029	0.3173	0.500
Wearing apparel, dressing and dyeing of fur	0.0012	0.347	0.243
Wholesale and retail trade - repairs	0.1258	0.4337	0.218
Wood and products of wood and cork	0.0026	0.2226	0.356

Note: A higher value of λ implies lower skill complementarity.

finance. These observations are in line with the model implication that highly-skilled workers sort into industries where skills are most substitutable and where their returns to human capital are highest (Proposition 1).

To provide additional evidence of this skill sorting, we use data from the CPS in 2000 to compute the average education achievement of workers in different industries. Figure 4 plots average education attainment (in years) versus the estimated skill complementarity in each industry: this exercise confirms that, on average, more educated individuals tend to sort in industries where skills are more substitutable.

Figure 4: Average education achievement versus estimated skill complementarity, by industry. Note: higher values on the x-axis correspond to higher skill complementarity.



Sources of persistence. The model features both an endogenous and an exogenous driver of intergenerational income persistence. The exogenous source is the persistence of heritable traits across generations; the endogenous source is the persistence of human capital attainment across generations. To assess their relative importance we shut down the persistence of heritable traits, holding their variance constant. In equilibrium the IGE is reduced to 0.324, suggesting that about $1/3$ of intergenerational persistence is due to exogenous transmission of traits, and $2/3$ is due to the endogenous persistence of human capital investments.

Assessing the effect of education spending. To provide external validation of the model we also assess its behavior. As a benchmark for comparison we use well-known findings by

Restuccia and Urrutia (2004). In particular, we examine an experiment involving increased public education expenditure. Restuccia and Urrutia's distinguish between college and lower education. We reproduce their experiments, in which education spending as a fraction of GDP is increased proportionally by 20% and the labor tax rate adjusts to finance new spending. That study estimates that the IGE would fall from a benchmark level of 0.4 to 0.36 when early education spending increases; in contrast, the IGE would exhibit no change if only college education spending were increased. The weighted average effect of these education experiments implies an IGE reduction from 0.4 to 0.378 in response to a 20% increase in education spending. For comparison we estimate the effect of a 20% increase in public education spending in our model, paid by higher labor income taxes: this would result in a reduction of the IGE from 0.47 to 0.443. The proportional reduction of the IGE in this experiment is roughly 6.1%, and very similar to the 5.8% proportional reduction estimated by Restuccia and Urrutia.

5 Counterfactual Experiments

To assess the explanatory power of the skill-substitutability mechanism we perform counterfactual experiments that answer the following question:

> *How different would intergenerational mobility in the U.S. be if its industrial composition was that of country 'X', but all other relevant features remained the same?*

There are two features of industrial composition that are easily observable across countries and can be varied in counterfactual experiments. In the first experiment we change the relative importance of an industry in the overall economy, as reflected in the share γ_n of final output paid to that industry. This aspect is important because if weight is shifted to any particular industry, overall skill substitutability will rise or fall depending on whether skills in that industry are relatively substitutable or complementary. The second experiment relates to the capital intensity in different industries. If skills are fairly substitutable within an industry, but the capital share is very large, then the effect on the overall complementarity will be limited. In contrast, if the capital share of output in that industry is small, then that industry will be more influential.

In addition to experiments that adjust output weights or capital shares, we also perform experiments in which we impose the tax and education policies observed in different countries on our US benchmark economy. These exercises serve two purposes: first, they provide a clear reference benchmark (based on the effects of policy differences on income mobility) which can be used to gauge the relative importance of skill complementarity for intergenerational mobility. In this sense, policy differences deliver a natural metric to quantify the importance of skill complementarity. Second, they also provide an upper bound on the indirect effect of skill complementarity differences. In other words, by superimposing a rough approximation of the tax and benefit system prevailing in a different country, we are able to measure how differences in public education spending — possibly due to underlying economic incentives — would affect intergenerational persistence in the US.

All of our experiments describe equilibrium outcomes, in which a new set of wages and other equilibrium objects are attained after the US equilibrium is perturbed by industrial and/or policy changes. One key constraint relates to the government budget identity. As overall skill substitutability alters incentives to invest in human capital, the equilibrium skill distribution will also change. This will impact government tax revenues, resulting in an unbalanced government budget constraint unless offsetting policy changes take place. We do not change the marginal tax rate because this would clearly alter the return to human capital (see Guvenen, Kuruscu, and Ozkan (2014)); for the same reason we do not want to alter the progressiveness of the tax system either. With this in mind, we allow the government budget constraint to be satisfied through a combination of changes in T and G, where the changes in T are restricted to maintain progressiveness at U.S. levels. This is equivalent to choosing T so that equation (31) is satisfied.

Table (5) reports the results of our experiments by country. For each experiment the results are divided into two parts, referring to the 'core' or 'core+5' samples. For each experiment we report the ratio of the standard deviation of counterfactual and observed IGEs, as well as the correlation between observed and predicted IGEs. The ratio of standard deviations indicates how much of the observed dispersion in IGEs is accounted for by the changes made in each experiment. The correlation is intended to show how well the predicted deviations from US mobility in each experiment align with observed deviations, regardless of the magnitude of deviations.

First, we focus on the experiment in which output shares are adjusted to reflect national data, holding all other exogenous features of the benchmark economy constant. There are some great successes in explaining observed IGEs as well as some failures. The experiment explains more than half of the difference in IGEs between Japan and the US, but at the same time explains little of the difference between Denmark and the US. The relative magnitudes of the predicted differences in IGE correspond to the relative magnitudes of observed differences as indicated by the sizeable correlations between data and experimental predictions, particularly for the core sample. In terms of explanatory power, output share differences account for 16% of IGE variation in the core sample and 18.5% in the core+5 sample.

Table 5: Experiments involve four alternative sets of changes. Each experiment measures the effect on the benchmark (US) IGE of adopting each different country's specific features. In column (1) we report the IGE estimate for each country, as in Corak (2006). In the other columns we report the counterfactual IGE obtained in the benchmark model after: (1) changing output shares γ_n; (2) changing both output and capital shares, γ_n and α_n; (3) changing observed policies and shares (adjust output and capital shares as well as education subsidization rate s and labor income tax rate τ); (4) changing observed policies only (adjust only public education and tax rates).

Country	IGE Estimate	Effect on benchmark (US) IGE of changing: Output Shares	Output and Capital Shares	Observed Policies and Shares	Observed Policies only
	(1)	(2)	(3)	(4)	(5)
USA	0.47	–	–	–	–
Core Sample					
Canada	0.19	0.410	0.398	0.398	0.466
Denmark	0.15	0.454	0.437	0.382	0.408
Finland	0.18	0.449	0.427	0.410	0.416
France	0.41	0.479	0.451	0.416	0.430
Norway	0.17	0.428	0.409	0.362	0.408
Sweden	0.27	0.448	0.451	0.428	0.410
Germany	0.32	0.443	0.401	0.393	0.422
UK	0.5	0.457	0.442	0.422	0.443
Correlation[b]	–	0.668	0.509	0.596	0.204
Relative S.D.[c]	–	0.160	0.171	0.174	0.160
Additional Sample					
Australia	0.26	0.438	0.411	0.422	0.472
Japan	0.34	0.399	0.392	0.405	0.476
Korea	0.25	0.426	0.413	0.437	0.482
Netherlands	0.23	0.453	0.459	0.430	0.449
Switzerland[a]	0.46	0.441	–	0.456	0.481
Correlation[b]	–	0.430	0.428	0.631	0.285
Relative S.D.[c]	–	0.185	0.205	0.220	0.260

Notes

a Capital share data are not available for Switzerland, thus US capital shares are maintained.

b The 'Correlation' raw reports the measure of association between observed and counterfactual IGEs for each given experiment.

c The 'Relative S.D.' raw reports how much of the observed IGE standard deviation is accounted for by the standard deviation of counterfactual IGEs in each given experiment.

When capital shares are adjusted, as well as output shares, the results generally improve. For example, this experiment explains twice as much of the difference between the US and Denmark, compared to when only output shares were adjusted, but total explanatory power for

Denmark remains low. Overall this experiment explains 17.1% of IGE variation in the core sample and 20.5% of IGE variation in the core+5 sample, both being slightly greater than the previous experiment.

In considering the effects of observed differences in education and fiscal policies, it is useful to compare the results of the experiment in which only observed policy changes are applied to the results of the experiments in which output and capital shares are changed. For the core sample, observed differences in public policy explain nearly an identical amount of the variation in IGE, as shown by the relative standard deviations. However, this variation generated by observed policy differences does not correspond nearly as closely to measured IGE variation as that generated by observed industrial composition differences. To see this note that the correlations between the fifth and first columns of Table (5) are about half as big as those between the third and first columns. Thus, taken together, we interpret these results as industrial composition doing a somewhat better job of explaining variation in IGEs than observed policy differences.

In the third experiment we change both technology shares and public policies to gauge their combined effects. As mentioned before, this can be viewed as an upper bound on the total effect of differences in skill complementarity on mobility that accounts for any induced variation in education subsidization motives. It is impossible to say how much of observed policy differences are actually due to differences in education spill-over across countries, but clearly an upper bound is all of it. One can see that explanatory power does modestly improve when policy differences are accounted for, particularly in the extended sample. Overall, we would conclude that skill complementarity can directly account for about 20% of international variation in mobility, and well above that level if allowing for indirect policy effects.

5.1 Discussion

The intergenerational persistence of economic status and the role of families in the transmission of privilege are the object of much debate by both academics and policy makers. Understanding the mechanics of this transmission is key to address questions about both economic inequality and efficiency. This paper focuses on supply-side incentives, which shape returns to parental human capital investments and affect both the observed persistence of earnings and the prevailing inequality. To what extent are observed patterns of intergenerational persistence due to underlying production incentives? How do families respond to such incentives? And what is the role of redistributive policies in accommodating efficient production?

To answer these questions we develop a tractable model which we estimate using a variety of micro data sources. Our findings suggest that long-standing differences in industry composition, and associated returns to human capital investment, may exert a significant influence on family investments in skills and induce different degrees of intergenerational persistence. Our estimates suggest that countries specializing in industries in which skills are more easily substitutable also exhibit significantly less intergenerational earnings mobility, as well as higher inequality of earnings. We argue that these differences are partly a reflection of production incentives, and we show that removing these incentives would considerably reduce observed cross-country differences in earnings mobility (IGE).

In the process we make some key assumptions about the economic environment. One particular assumption warrants further discussion. Our theory implicitly assumes that causality runs from industrial composition to the IGE. Effectively, we take the persistent cross-country differences in industrial composition as exogenously given. A much more ambitious, and demanding, approach would allow for the joint determination of technology adoption and human capital investments. However, even a much richer theory would ultimately need to appeal to some primitive differences across economies (deep-rooted sources of comparative advantage or historical events) in order to rationalize the persistent and slow moving differences in industry composition, skills and employment that we observe in the data.

6 Conclusions

We explore the hypothesis that different production arrangements help shape the transmission of economic advantage, and explain the large discrepancies in earnings mobility observed across countries.

First, we illustrate how the intensity of strategic complementarity in parental skill investments may lead to more or less earnings' mobility. In particular, we show that stronger skill substitutability implies a higher degree of intergenerational persistence and may be associated with less progressive public policies that equalize skills in the workforce. To provide evidence in support of this hypothesis we study geographical variation in income persistence, linking skill substitutability in production and intergenerational mobility at the country level. We find a significant cross-country association between different measures of skill substitutability and IGEs.

Next, to explore the origins of these statistical associations we develop and estimate a structural equilibrium model, and we use it to broadly quantify the importance of skill substitutability for cross-country differences in intergenerational mobility. The model explicitly allows for: (i) an exogenous persistent process for heritable traits (skill endowments); (ii) endogenously persistent investments in human capital; and (iii) an equilibrium distribution of human capital attainments with associated market-clearing wages. The production side of the economy consists of many industries, and for each such industry we estimate a parameter summarizing the degree of workers' skill substitutability. We derive a theoretical result that characterizes equilibrium worker-industry matching, and use it to solve for the model equilibrium.

We perform various counterfactual experiments, which involve re-weighting to reflect the industry composition of different countries. These experiments indicate that between 15 and 20% of international variation in intergenerational mobility can be directly accounted for by cross-country differences in skill substitutability in production.

Finally, we discuss the relationship between public education policies and skill complementarity. The model implies that optimal education subsidies should be higher in countries where skill substitutability in production is weakest. In fact, these countries should exhibit a mix of more progressive policies and higher earnings' mobility. This observation suggests the presence of an indirect effect through which technology may affect intergenerational mobility, namely by changing social incentives to adopt policies that equalize skills in the working population. When accounting for these indirect effects the model explains a larger share of the cross-country variation in earnings mobility, and provides a way to rationalize the observation of significant and persistent geographic differences in both economic mobility and progressiveness of government policies.

References

ABBOTT, B. (2015): "The Effect of Parental Composition on Investments in Children," Discussion paper, Queen's University.

ABBOTT, B., G. GALLIPOLI, C. MEGHIR, AND G. VIOLANTE (2013): "Education Policy and Intergenerational Transfers in Equilibrium," Discussion paper, National Bureau of Economic Research.

ARROW, K. J. (1962): "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, pp. 155–173.

BARRO, R. J., AND J. W. LEE (1994): "Dataset for a panel of 138 countries," Discussion paper, NBER, Available at http://www.nber.org/pub/barro.lee/.

——— (2013): "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 104, 184–198.

BECKER, G. S. (1981): "A Treatise on the Family," *NBER Books*.

BECKER, G. S., AND N. TOMES (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), S1–39.

BLACK, S. E., AND P. J. DEVEREUX (2011): "Recent Developments in Intergenerational Mobility," *Handbook of Labor Economics*, 4, 1487–1541.

BLANDEN, J. (2009): "How Much Can We Learn From International Comparisons of Intergenerational Mobility?," Discussion paper, Centre for the Economics of Education, London School of Economics and Political Science.

——— (2013): "Cross-Country Rankings in Intergenerational Mobility: a Comparison of Approaches from Economics and Sociology," *Journal of Economic Surveys*, 27(1), 38–73.

BOMBARDINI, M., G. GALLIPOLI, AND G. PUPATO (2012): "Skill Dispersion and Trade Flows," *American Economic Review*, 102(5), 2327–48.

——— (2014): "Unobservable Skill Dispersion and Comparative Advantage," *Journal of International Economics*, 92(2), 317–329.

BOWLUS, A. J., AND J.-M. ROBIN (2012): "An International Comparison of Lifetime Inequality: How Continental Europe Resembles North America," *Journal of the European Economic Association*, 10(6), 1236–1262.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014): "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 129(4), 1553–1623.

CORAK, M. (2006): "Do Poor Children Become Poor Adults? Lessons for Public Policy from a Cross-Country Comparison of Generational Earnings Mobility," *Dynamics of Inequality and Poverty*, 13, 143–188.

——— (2013): "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *The Journal of Economic Perspectives*, pp. 79–102.

CORTES, G. M., AND G. GALLIPOLI (2015): "The Costs of Occupational Mobility: An Aggregate Analysis," *Working Paper (SSRN 2717671)*.

CUNHA, F., J. J. HECKMAN, AND S. M. SCHENNACH (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931.

DEL BOCA, D., C. FLINN, AND M. WISWALL (2014): "Household Choices and Child Development," *Review of Economic Studies*, 81(1), 137–185.

GOLDIN, C., AND L. F. KATZ (1995): "The Decline of Non-Competing Groups: Changes in the Premium to Education, 1890 to 1940," Working Paper 5202, National Bureau of Economic Research.

GUVENEN, F., B. KURUSCU, AND S. OZKAN (2014): "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *The Review of Economic Studies*, 81(2), 818–850.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2010): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

ICHINO, A., L. KARABARBOUNIS, AND E. MORETTI (2010): "The Political Economy of Intergenerational Income Mobility," *Economic Inquiry*, 49(1), 47–69.

JANTTI, M., B. BRATSBERG, K. RØED, O. RAAUM, R. NAYLOR, E. OSTERBACKA, A. BJORKLUND, AND T. ERIKSSON (2006): "American Exceptionalism in a New Light:

a Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States," *IZA Discussion Paper*.

KRUEGER, A. (2012): "The Rise and Consequences of Inequality," Discussion paper, Center for American Progress, Presentation made on January 12th. Available at http://www. americanprogress. org/events/2012/01/12/17181/the-rise-and-consequences-of-inequality.

LEE, S. Y. T., AND A. SESHADRI (2014): "On the Intergenerational Transmission of Economic Status," Discussion paper, Mimeo, University of Wisconsin, Madison.

MCDANIEL, C. (2014): "Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD 1950-2003," Mimeo, Arizona State University.

MORETTI, E. (2004): "Workers Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *The American Economic Review*, 94(3), 656–690.

PYCIA, M. (2012): "Stability and Preference Alignment in Matching and Coalition Formation," *Econometrica*, 80(1), 323–362.

RESTUCCIA, D., AND C. URRUTIA (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education," *The American Economic Review*, 94(5), 1354–1378.

ROMER, P. M. (1986): "Increasing Returns and Long-Run Growth," *The Journal of Political Economy*, pp. 1002–1037.

SOLON, G. (2004): "A Model of Intergenerational Mobility Variation over Time and Place," in *Generational Income Mobility in North America and Europe*, ed. by M. Corak, pp. 38–47. Cambridge University Press, Cambridge.

STORESLETTEN, K., C. I. TELMER, AND A. YARON (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51, 609–33.

A Appendix

A.1 Cross-country industry composition: STAN data

We use data provided by the OECD to approximate the relative importance of different industries within each of the countries considered in the reduced form analysis. The data is from the structural STAN databases and provides information about the value added of each individual industry in different years. The data is disaggregated at the 2-digit level and covers 30 industry groups, which are listed in the following table.

Table 6: This table reports the industries used in our analysis. The first column indicates the ISIC code corresponding to each industry (or subset of codes, e.g. C01T05 means from C01 to C05.)

ISIC code	Industry
C01T05	Agriculture, hunting, forestry and fishing
C10T14	Mining
C15T16	Food products, beverages and tobacco
C17	Textiles
C18	Wearing apparel, dressing and dyeing of fur
C19	Leather, leather products and footwear
C20	Wood and products of wood and cork
C21	Pulp, paper and paper products
C22	Printing and publishing
C23T25	Chemical, rubber, plastics and fuel products
C26	Other non-metallic mineral products
C27T28	Basic metals and fabricated metal products
C29	Machinery and equipment, n.e.c.
C30T33	Electrical and optical equipment
C34	Motor vehicles, trailers, and semi-trailers
C35	Other transport equipment
C36T37	Manufacturing n.e.c. and recycling
C40T41	Electricity, gas and water supply
C45	Construction
C50T52	Wholesale and retail trade - repairs
C55	Hotels and restaurants
C60T63	Transport and storage
C64	Post and telecommunications
C65T67	Financial intermediation
C70	Real estate activities
C71T74	Renting of mach. and equip. - other business activities
C75	Public admin. and defence - compulsory social security
C80	Education
C85	Health and social work
C90T93	Other community, social and personal services

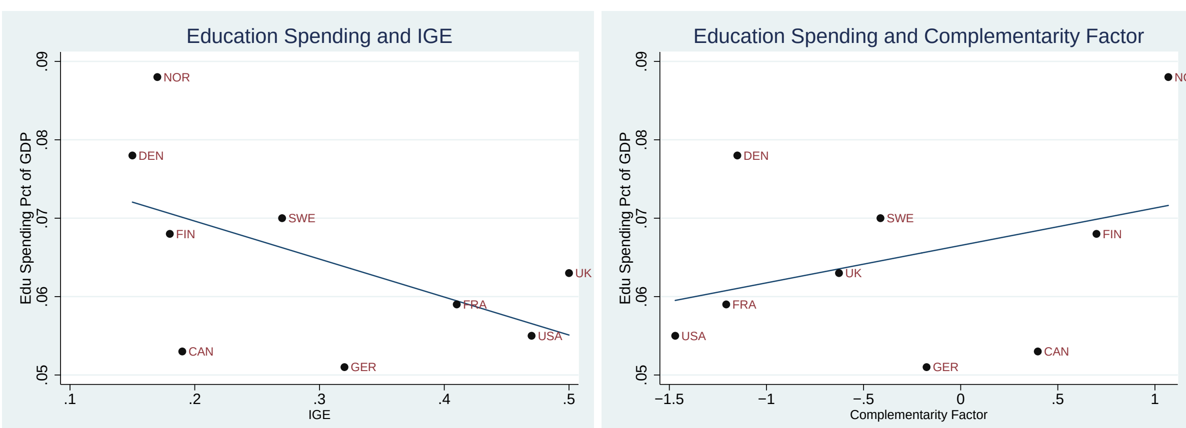
A.2 Taxes, Education Attainment, and Public Education Spending

Table (7) reports both current (2010) and historic (1960-64) public education expenditure levels, as well as labor income tax burdens for the year 2010. Current information is primarily based on OECD data, whereas the historical education spending data is based on the work of Barro and Lee (1994).

In addition to their use in some of our counterfactual experiments, these data also help to illustrate a few interesting relationships. In particular, our simple model in Section 2 has implications for the relationship among public education spending, IGE and skill complementarity. The literature (for an example, see Table 8 in Blanden, 2013), has shown that public education spending as a fraction of GDP is higher in countries that have lower IGEs.

Combining Tables (7) and (1) one can see that this is also true in the data we have presented. The left panel of Figure (5) shows exactly this by plotting the percentage of GDP spent on education by governments against the IGEs, for our core sample. Less obvious is whether this relationship can possibly be attributed to skill complementarity. In the right panel we explore this question by also plotting public education spending against the estimated complementarity proxy (the O*Net common factor) for the core sample.[21] As the theory predicts, higher education spending is observed in countries where skill complementarity is greater. However, we note that the slope of the regression line is not significant at conventional levels, thus more work would be needed before drawing firm conclusions.

Figure 5: Public education spending, IGE and skill complementarity common factor proxy.



Another prediction of our model (apparent in equation (7)) is that, *ceteris paribus*, less

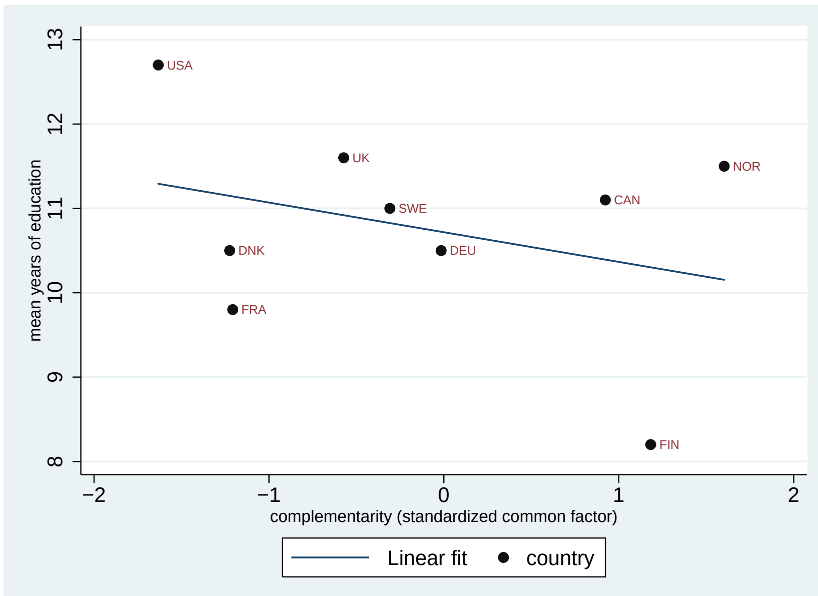
[21]In the extended sample Korea and Japan are large outliers in education expenditure, thus we only plot the core sample.

Table 7: The second column of this table reports the share of GDP spent by all levels of government on education in year 2010, as per the OECD document "Education at a Glance 2013." *German data taken from World Bank data. The third column also reports the percent of GDP that is public education spending, but for an earlier period 1960-64, as reported in Barro and Lee (1994). The fourth column shows average education attainment (in years of schooling) for 2000, as measured by Barro and Lee (2013). The fifth column reports national average labor income tax rates in 2010, also reported by the OECD. The sixth column reports the average estimated complementarity (O*Net common factor) by country, i.e. the measure used in panels (e) and (f) of Figure 3.

Country	Public Education Spending as % of GDP (2010)	Public Education Spending as % of GDP (1960-64)	Average Education Attainment (2000)	Average Tax Rate	O*Net Common Factor
Core Sample					
United States	5.5%	6.37%	12.7	29.6%	-1.470
Canada	5.3%	7.70%	11.1	30.8%	0.397
Denmark	8.8%	7.29%	10.5	38.6%	-1.150
Finland	6.8%	5.79%	8.2	42.5%	0.699
France	5.9%	4.42%	9.8	50.2%	-1.207
Norway	8.8%	6.71%	11.5	37.6%	1.069
Sweden	7.0%	7.23%	11.0	42.8%	-0.413
Germany	5.1%*	4.26%	10.5	49.7%	-0.176
UK	6.3%	5.14%	11.6	32.3%	-0.627
Additional sample					
Australia	5.2%	4.97%	11.9	27.2%	0.345
Japan	3.8%	4.48%	10.8	31.2%	1.253
Korea	4.9%	4.14%	10.6	21.0%	2.119
Netherlands	6.0%	7.48%	10.8	38.6%	-0.404
Switzerland	5.2%	4.48%	11.4	21.5%	-0.572

skill complementarity (larger λ) implies greater education attainment. Of course, as the previous graph has shown the *ceteris paribus* argument does not hold because public education spending tends to be greater where complementarity is stronger. It is nonetheless interesting to examine the relationship between education attainment and skill complementarity. To gauge the intensity of this relationship we use data on education achievement in different countries, for the year 2000, provided by Barro and Lee (2013). As Figure 6 shows the negative relationship implied by the theory does appear in the data, although it is not estimated precisely.[22]

Figure 6: Average education attainment and skill complementarity, by country.



A.3 Cross-Country Patterns in Industry Wage and Income Dispersion

As the degree of substitutability of skills is not directly observable, we pursue different ways to rank industries in terms of their ability to substitute across workers' skills. One of them is to exploit a theoretical result linking the degree of complementarity to the measured dispersion of raw, and residual, wages within industries. In a setting with labor market frictions, Bombardini, Gallipoli, and Pupato (2012) establish that wage dispersion within industries increases in the degree of skill substitutability when some skills are unobservable. Sectors with higher complementarity are characterized by a more compressed wage distribution because, for example, workers with higher-than-average skills contribute relatively less to surplus, a fact reflected in

[22]Results are similar if one uses the Barro-Lee data for the year 2005.

their wage. Despite the different model setting, a similar result holds also in the context of our analysis: the proof of Proposition (1) shows that the ratio of the marginal products of high to low skilled workers will always be larger in industries where skills are more substitutable. In other words, given two workers with different skills, the difference in their wages will increase in the degree of skill substitutability in production. This implication of the model is also apparent when looking at equation (28), which links industry-specific skill substitutability to raw wage dispersion: differences in wage dispersion across sectors are partly due to differences in the way skills are aggregated and one should observe a positive correlation in industry-specific wage dispersion across different countries. Moreover, based on the proof of Proposition (1), this cross-country correlation at the industry level should continue to hold even after purging out some workers' heterogeneity, after controlling for their observable characteristics.

We set out to investigate this hypothesis and document that the ranking of within-industry wage dispersion follows a consistent pattern across countries by using the Luxembourg Income Study Database (LIS). The LIS provides a set of cross-sectional datasets describing household and individual income and other characteristics for a large number of countries and years. These datasets have been harmonized (to the extent possible) to make variables directly comparable across countries and years.

Unfortunately, while the variables provided in the LIS are comparable across datasets, many variables are only available for a limited number of datasets. In particular, there is insufficient wage-by-industry data for Australia, Canada, Japan, Korea, Norway, Sweden, and Switzerland. However, good data exists over multiple years for the US, Germany, and Ireland. A handful of EU countries have a single year with sufficient data. The UK and France[23] lack wage data, but have data on total labour income. Therefore, we present two comparisons of within-industry wage dispersion: one based on total labour income, and the other based on hourly wage.

For income statistics, our sample includes all individuals between the ages of 16 and 65 with non-zero wages. Individuals are weighted by the population weight provided by the LIS. For wage statistics, we are able to restrict the sample to individuals who are employed in private industry, and we weigh individuals by their average (weekly) hours worked as well as their population weight. We are unable to consistently identify self-employed individuals, so they remain in both samples.

Our 30 industry classification system is based on 2-digit ISIC3 industry codes. Many

[23]Income data for France is not perfectly comparable to the other countries in the sample because it is reported after certain deductions are made. However, we feel that even imperfect data can be informative so we include it nonetheless.

datasets in the LIS include 2-digit ISIC industry codes or a compatible classification as a part of their labour statistics, making conversion straightforward. The US and France use unique classification systems. For these countries we constructed our own crosswalks based on the documented descriptions of the industry codes. We rank industries according to the standard deviation of log labour income and log wage (from highest to lowest), both with and without controls. This results in four different sets of ranks overall. We use the same set of controls for both labour and wage ranks. Namely: industry (by our 30 industry classification), education (a three category classification), age (with squared and cubed terms), sex, and region (state or province, depending on the country).

We calculate the four rankings for the US and for a set of EU countries, according to data availability. Estimating standard deviations for each industry requires a large number of individuals in every industry. The LIS datasets sometimes have a very small number of individuals in particular industries. This leads to less reliable estimates of ranks. However, ranks are unlikely to shift substantially for each country over a short span of time. Therefore we also calculate pooled ranks for a few countries where this is possible. More specifically, we pool individuals in every viable dataset between 1999 and 2014 for the given country and add year dummies to the set of controls.

Figure (7) shows the relationship between ranks for US and Germany based on pooled samples. There is a strong correlation between ranks, regardless of which of the four statistics we use. Rankings based on income dispersion are almost identical between the two countries. The correlation between wage dispersion ranks is comparatively weaker but remains quite strong and significant. This is partly because the sample sizes are quite a bit smaller, which adds noise to the rank measures. In both cases, adding controls reduces the strength of correlation, indicating that common demographic patterns explain a small part of the correspondence between rankings. However, the fact that even the ranks calculated from wages with controls show a significant positive correlation provides robust evidence that there is some unobserved feature of industry structure which characterizes each industry and is common across countries, influencing wage dispersion.

The widespread nature of this pattern is documented in Table (8). In this Table we replicate the exercise illustrated in Figure (7) for a larger set of countries. Ranks are computed for each dataset (or pooled dataset), and then each rank is regressed individually on the corresponding ranks from the pooled US sample.[24] Table (8) reports the resulting slope coefficients and standard deviations for OLS regressions with robust standard errors.

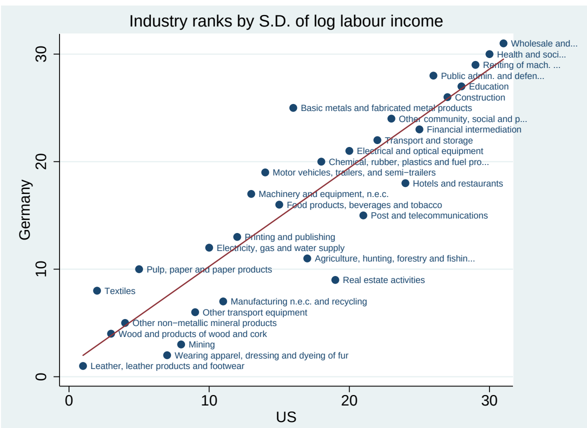
The results in Table (8) show that the industry ranks calculated for all countries in the sam-

[24]We drop the "residual" industry category, as it is sometimes an outlier and is not particularly informative.

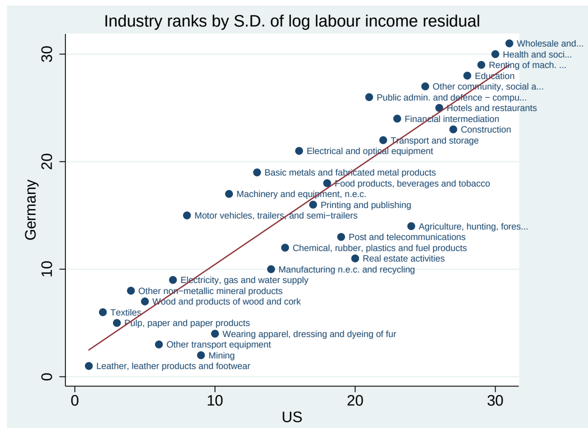
ple are highly correlated. As in the case of Germany, the correlation for ranks of labor income dispersion is stronger than for wage dispersion. Yet, in the whole sample, controlling for demographic characteristics has a small effect, and does not always decrease the strength of the relationship. This suggests that the main source of these correlations is some other unobserved industry specific effect. Again, this evidence supports the argument that the international pattern of within-industry wage dispersion must be due to some aspect of industry structure.

Our hypothesis is that this pattern reflects different degrees of skill complementarity in the production process. While countries may be using slightly different production technologies to produce similar goods, these technologies may all share some features which broadly shape the organizational structure and types of labour employed: low skill-complementarity technologies allow firms to hire a mix of high and low skill workers, while high skill-complementarity technologies encourage a more homogenous workforce, as shown in Bombardini, Gallipoli and Pupato (2012). This is consistent with the observed covariation in the international patterns of wage dispersion by industry.

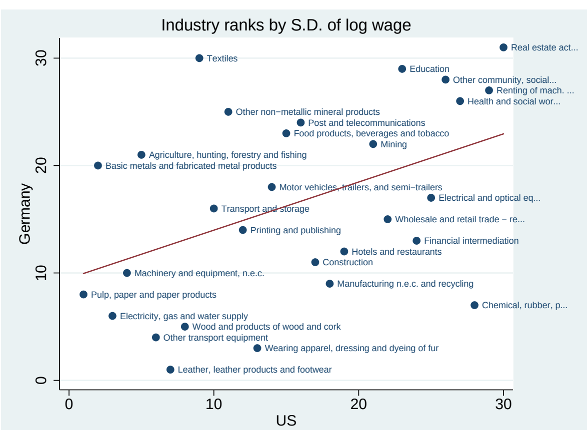
Figure 7: Comparison of US and German industry ranks according to different dispersion statistics. "Residual" refers to the remaining unexplained variation after controlling for industry, education, age, sex, region (state or province), and year.



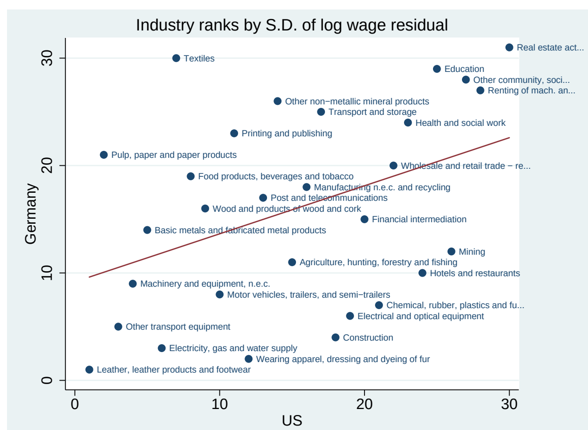
(a) Log labour income dispersion ranks



(b) Log labour income residual dispersion ranks



(c) Log wage dispersion ranks



(d) Log wage residual dispersion ranks

Table 8: This table summarizes the correspondence of within-industry earnings and wage dispersion patterns across countries, based on comparable LIS data. Values are calculated in two steps: first we calculate the standard deviation of each (logged) variable for each industry in a given country, and rank industries in that country from highest to lowest dispersion; second, the ranks of each country and dispersion statistic are regressed independently on the corresponding ranks for the US using OLS with robust standard errors. The table reports the slope coefficient associated with each regression. Standard errors are in brackets. Missing values occur where the LIS has insufficient data on wages.

Industry S.D. rank, regression slope coefficients

	Dependant variable: industry S.D. rank, US data (pooled 2004,2007,2010)			
	log labour income	log labour income residual	log wage	log wage residual
	(1)	(2)	(3)	(4)
Regressor variable				
Pooled years				
Germany 2004,2007,2010	0.92	0.88	0.45	0.45
	(0.00)	(0.00)	(0.03)	(0.03)
Ireland 2000,2004,2007,2010	0.90	0.93	0.29	0.58
	(0.00)	(0.00)	(0.03)	(0.02)
UK 1999,2004,2007,2010	0.90	0.89		
	(0.00)	(0.01)		
Individual years				
USA 2010	1.00	0.99	0.86	0.93
	(0.00)	(0.00)	(0.01)	(0.00)
Germany 2010	0.93	0.89	0.47	0.61
	(0.00)	(0.00)	(0.02)	(0.02)
Ireland 2010	0.91	0.88	0.48	0.59
	(0.01)	(0.01)	(0.03)	(0.02)
UK 2010	0.56	0.52		
	(0.02)	(0.02)		
France 2005	0.61	0.67		
	(0.03)	(0.02)		
Austria 2004	0.79	0.76		
	(0.01)	(0.01)		
Belgium 2000	0.83	0.82	0.42	0.44
	(0.01)	(0.01)	(0.02)	(0.04)
Spain 2000	0.81	0.80	0.26	0.35
	(0.01)	(0.01)	(0.03)	(0.04)
Finland 2007	0.88	0.87	0.47	0.56
	(0.01)	(0.01)	(0.03)	(0.02)
Greece 2004	0.79	0.82		
	(0.01)	(0.01)		