Skill Complementarity and the Geography of Intergenerational Mobility

Preliminary and Incomplete Draft

Brant Abbott

Yale University

Giovanni Gallipoli

University of British Columbia

September 26, 2014

Abstract

We develop and estimate an equilibrium model of geographic variation in the intergenerational earnings elasticity (IGE). The theory extends the Becker-Tomes model, introducing a production sector in which human capital inputs are strategic complements. We show that the equilibrium return to human capital investments is lower in places where strategic complementarity is more intense, and that this is associated to less intergenerational persistence (lower IGEs). Furthermore, optimal education policies are more progressive where these complementarities are stronger, leading to a negative correlation between progressive public policy and IGEs. Using microdata we construct various location-specific measures of skill complementarity and document that the patterns of geographic variation in IGEs are consistent with our hypothesis. Quantitatively, geographic differences in skill complementarity account for up to 1/5 of cross-country variation in intergenerational earnings persistence. Governments in countries where prominent industries exhibit greater skill complementarities tend to spend larger fractions of GDP on public education, suggesting that underlying technology differences may indirectly explain an even larger proportion of cross-country IGE variation.

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). More recent work has measured intergenerational earnings mobility across regions of the United States (Chetty, Hendren, Kline, and Saez, 2014), finding large and persistent differences. This body of research has also documented interesting correlations between IGEs and measures of public education spending, inequality and returns to human capital investments (Blanden, 2009). Yet, little is known about what drives these correlations and whether they are useful to understand 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 both variation in 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 strategic complementarities in the way human capital is used in production directly affect the intergenerational persistence of earnings. We also show that differences in strategic complementarity result in differences in the desirability of progressive public education policies, highlighting an additional (indirect) channel through which technology may affect IGEs.

When we analyze 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. Our results suggest that differences in the degree of skill-complementarity can account for about 20% of international variation in IGEs. For comparison, observed variation in the generosity of public policies can explain about 25% of the 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. These 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, workers' productivity is less dependent on co-workers' skills.

[1] Quantitative studies of intergenerational persistence have focused on the 'aggregate' mobility rate of the United States (e.g. Restuccia and Urrutia, 2004; Lee and Seshadri, 2014).

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 overall skill substitutability across countries. Comparative advantage in certain industries, and technologies, may therefore affect human capital investments, government policies, and earnings 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 which equalize skills will be more desirable. We present evidence that these relationships cannot be rejected in cross-sectional data from a variety of developed countries.

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

The more substitutable are skills, the greater the effect one's *own* skill attainment has on wages. Thus, the human capital investment made by a parent will have a greater effect on their child's future earnings if they live in a country where skill substitutability is higher. Moreover, the greater returns to human capital investment in those countries where skill substitutability is higher will induce larger human capital investments.

Variation in the degree of skill substitutability may also have implications for the optimal progressivity of education and tax policies, and thus have indirect effects on intergenerational mobility. This is because the degree of strategic complementarity in human capital investments increases as skills become more complementary, and an increase in the aggregate stock of human capital has a larger effect on aggregate efficiency. This 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

[2]Moretti (2004) shows that the spillover is larger between similar industries than otherwise. His industry decomposition is relatively fine, and hence the 'similar' industries in his data mostly fall within the same coarse industry at our level of aggregation.

and intergenerational mobility endogenously arises in such framework. This also provides an additional channel through which skill substitutability can affect intergenerational mobility.

Finally, a positive correlation between income inequality and economic mobility arises in this model, as the degree of income inequality is directly related to the substitutability of skills in production.

We begin the paper with an analytical example which illustrates the mechanism. In this simple setting we consider only two periods and we let parents' human capital levels be exogenous endowments. There are no heritable skills, thus a parent only affects the child's outcomes by investing in her human capital. Altruism motivates parents to do so; however the relationship between the human capital that parents bestow upon their children and children's earnings depends on the skill substitutability parameter in the aggregate production function. It turns out that these simplifications lead to the stylized result that the IGE equals 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 in skill investment.

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). Using these measures we compute the overall skill substitutability within different countries, weighing each industry by its relative size obtained from OECD-STAN data. Then we examine the cross-country relationship between aggregate skill substitutability and estimates of IGEs taken from the literature. The results are very interesting. Using a variety of skill substitutability measures we consistently find a robust negative correlation between the IGE and the degree of skill substitutability prevailing in a country, exactly as the analytical example suggests.

We also replicate this reduced-form analysis at the level of US commuting zones, using measures of intergenerational income persistence provided by Chetty, Hendren, Kline, and Saez (2014). For this exercise we construct measures of skill substitutability in production for each commuting zone using CPS data; then we examine how geographic variation in mobility correlates with proxies of skill substitutability in production across regions. As for the cross-country analysis, we find a significant relationship between measures of substitutability and levels of income persistence.

This empirical evidence suggests that some relationship exists 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 model entailing a steady-state overlapping-generations economy with an endogenous skill distribution and multiple sectors. Workers in this economy are free to select into different industries, each with an associated level of skill substitutability. Total production of the final consumption good in this economy is a Cobb-Douglas aggregate of the goods produced in the various sectors. This model is solved and estimated using US data. Then, to assess the role of technological differences, we run counterfactual experiments in which we change the industry composition to generate the degree of skill substitutability observed in other countries: this corresponds to re-weighing industries so that aggregate industrial composition reflects that of a different country, where the intergenerational income elasticity is also known. Crucially, public education policy, marginal tax rates, and the progressivity of the tax system are held constant at US levels in order to isolate the direct effect of technology. Our results indicate that between 15% and 20% of international variation in IGEs can be explained by differences in human capital returns induced by variation in overall skill substitutability.

Lastly, we explore the possibility that strategic complementarities may affect IGEs through the additional channel of optimal public policies. We solve a set of constrained social planning problems for a select set of countries, and obtain the degree of skill subsidization that maximizes welfare conditional on the prevalent technology structure. When we allow for this indirect effect of technology on policy the model can account for up to 1/3 of observed international variation in IGEs. A caveat is in order: one should not draw conclusions about how much of the actual variation in IGEs is a result of 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 some evidence that countries characterized by lower levels of skill substitutability in production tend on average to 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 that setting. Section 3 presents evidence of the reduced-form relationship between mobility and skill substitutability, across US regions as well as across countries. Section 4 describes the richer structural model and how we estimate it. Section 5 overviews the experiments and findings, while Section 6 introduce 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 relation 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 some analytical tractability would be lost. In contrast, the two-period model yields clear predictions that are easily interpretable.

2.1 Setup

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

The labor market is competitive. An agent with human capital h is paid a wage $w(h)$ per unit of human capital supplied. The production technology is a CES aggregate of the human capital of all workers:

$$y = \left(\int h_i^{\lambda} di \right)^{\frac{1}{\lambda}}. \tag{1}$$

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 skills implies a higher 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, which equals the marginal product of human capital:

$$w(h) = y^{1-\lambda} h^{\lambda-1}. \tag{2}$$

Other workers' productivities influence a person's wages through y, but the influence of y diminishes as λ approaches unity.

A 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

$$V_p(h_p) = \max_{c_p, h_c} \left\{ u(c_p) + \beta V_c(h_c) \mid c_p + h_c = w(h_p) \cdot h_p) \right\}. \tag{3}$$

The constraint in the maximization problem is the parent's budget constraint. Their own consumption and their investments in their child must be paid for out of labor income. Implicit in this restriction is that parents cannot borrow against the child's future income in order to finance human capital.

2.2 Implications for Earnings Mobility

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

$$V_c(h_c) = \max_{c_c} \left\{ u(c_c) \mid c_c = w(h_c) \cdot h_c) \right\}. \tag{4}$$

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

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

$$h_c^{1-\lambda} = \frac{u'(c_c)}{u'(c_p)} \beta \lambda y_c^{1-\lambda}. \tag{5}$$

This equation obviates the complexity of a parents problem. Their choice of human capital investment in their child is related to both child's and own marginal utility of consumption, as well as aggregate economic conditions.

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

$$h_c = \frac{\beta \lambda}{1 + \beta \lambda} y_p^{1-\lambda} h_p^{\lambda}. \tag{6}$$

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 others' skills on one's own human capital increases as skills become more complementary in production.

This expression can also be written in terms of income. Let $x = w(h)h$ be household income. Then equation 6 can be rearranged into:

$$x_c = \left(\frac{\beta \lambda}{1 + \beta \lambda} \right)^{\lambda} y_c^{1-\lambda} x_p^{\lambda}. \tag{7}$$

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

2.3 Optimal Education Policy

Institutional variation in government support for human capital investments is often associated to higher intergenerational mobility. In cross-country data the level of public support appears to be negatively related to the intergenerational income elasticity, indicating that heavier public education subsidies are associated with greater economic mobility. What is unclear is whether there is a causal relationship, or this coincidence is the 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 λ is negatively related to the optimal subsidy rate and the optimal level of public education support is positively related to intergenerational mobility.

The constrained social planning problem we consider involves maximizing ex-ante social welfare 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 prior to learning the relative advantage of the parent. Hence, the planning problem is as follows:

$$\begin{aligned} &\max_{s,\tau} \int V_p(h_p; s, \tau) dF(h_p) \\ &\qquad s.t. \\ &c_p + (1-s)h_c = (1-\tau) y_p^{1-\lambda} h_p^{\lambda} \\ &c_c = y_c^{1-\lambda} h_c^{\lambda} \\ &h_c = \tfrac{1-\tau}{1-s} \tfrac{\beta\lambda}{1+\beta\lambda} y_p^{1-\lambda} h_p^{\lambda} \\ &\tau y_p = s \int h_c di \\ &y_c = (\textstyle\int {h_c}^{\lambda} di)^{\frac{1}{\lambda}} \end{aligned} \tag{8}$$

Crucially, the third constraint imposes that household optimization holds, i.e. this is a Ramsey planning problem. The fourth constraint imposes that a government budget constraint must be satisfied 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 human capital investments. After some algebra, the planner's optimal policies reduce to,

$$\begin{aligned} s* &= 1-\lambda \\ \tau* &= (1-\lambda)\tfrac{\beta}{1+\beta}. \end{aligned} \tag{9}$$

Clearly, as the degree of strategic complementarity rises the social planner finds interventions more desirable. In this simplistic case the optimal subsidy is in fact exactly the degree of skill substitutability in production.

Although substantial interventions would occur if strategic complementarity is strong enough, this would not actually affect the intergenerational earnings elasticity. This can easily be seen by writing a version of equation 7 that would hold under these policies:

$$x_c = \left[\frac{\beta\lambda(1+\beta)}{\beta(1+\beta\lambda)}\right]^{\lambda} y_c^{1-\lambda} x_p^{\lambda}. \tag{10}$$

The intercept of a generational log-earnings regression would be affected, but not the slope. However, this independence between policy and the intergenerational elasticity is not a general result. If public investment are lump-sum rather than proportional, which may be a more realistic representation of public primary/secondary schooling systems, intergenerational mobility will be affected.

Augmenting our setup so that $h_c = m + S$, we would have the following expression for a child's human capital attainment:

$$h_c = \max\left\{S, (1-\tau)\frac{\beta\lambda}{1+\beta\lambda} y_p^{1-\lambda} h_p^{\lambda}\right\}. \tag{11}$$

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 it to the richer numerical analysis below.

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 in different locations, which we then relate to local measures of inter-generational earnings persistence. The analysis is performed at different levels of aggregation: first, we look at cross-country variation; then, we focus on the U.S. and consider differences across commuting zones[3]

3.1 Cross-Country Differences

The hypothesis that intergenerational persistence within a country may be partly related to that country's production arrangements is hard to test directly. The lack of accurate, and comparable, data on income 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 intergenerational persistence of income and economic status.

To approximate 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 some interesting robustness checks of our reduced-form results.

Furthermore, for each country in our samples, we construct skill substitutability proxies which subsume differences in their production structure, as reflected 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. Next, we briefly describe how different proxies are constructed.

Industry measures of skill substitutability. We use two different approaches to measuring skill substitutability at the industry level[4]. The first relies on model restrictions, which suggest that industries in which skill substitutability is relatively stronger are, ceteris paribus, characterized by higher wage dispersion. This implication follows from equation 6. The wage dispersion proxies are obtained using US data from the CPS in 2000. We present results using the standard deviation and the coefficient of variation as measures of wage dispersion, but we

[3]Among the many differences inherent in these two levels of analysis are (1) worker migration is less of a concern at the international level than the US commuting zone level due to the barriers described in Klein and Ventura (2009); and (2) differences in the skill complementarity of particular industries across places are likely to be less severe within the US than across countries.

[4]For a similar approach see Bombardini, Gallipoli, and Pupato (2012; 2014).

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 regressing measures of individual wages on a set of observable characteristics such as age, education, and gender. 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 31 different industries. To obtain industry weights we take an average of each industry's value-added share over the five years' interval between 2001 and 2005, as recorded in the STAN. Finally, using these weights we build measures of wage dispersion for each country.

To assess the robustness of results we also use a different set of skill substitutability proxies, which do not rely on economic theory but rather exploit direct measures of skill substitutability obtained from the O*NET database. The O*NET reports information on many occupation-specific requirements, including measures of skill substitutability in production. 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 which 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 which 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 three questions ranges between 1 and 5. The second question is slightly more nuanced, asking workers: "How 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 is aimed to measure directly the amount of co-workers' output which depends on each worker input.

In summary, we end up with two subsets of substitutability measures, each consisting of four different proxies. In the O*NET subset we measure the importance of team membership, responsibility for others' output and independence in production. In the subset of wage dispersion measures we include the standard deviation and the coefficient of variation for both raw and residual wages. 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. Each proxy subset contains four separate measures of skill substitutability in production. We first use each of them separately; then, we perform factor analysis to identify the principal component driving the measures within each subset. This gives us a unique proxy for skill substitutability in different industries which should in principle be more precise, as it is generated by aggregating the noisy information of all proxies in a given subset. An additional restriction, satisfied by all the measures we use, is that the proxies within each subset must positively correlate across different occupations and industries, indicating that they similarly co-vary with skill substitutability. This restriction also makes the interpretation of the common components relatively straightforward.

Country measures of 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 partly control for this problem, we construct different samples of countries for which we observe IGEs. 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 estimates of the IGE exist for them and because many of these estimates of the IGE are relatively comparable, and reliable. Multiple measurements allow researchers to form a fairly good idea of the value of IGE in these countries. An additional advantage of starting from this small set of nine countries is that Corak provides a set of 'low end' and 'high end' estimates of the IGE for them. We use this extra information to verify the robustness of the cross-country results.

To gauge the robustness of our findings we also extend the core sample to 5 more countries, namely Australia, Japan, Korea, Netherlands and Switzerland. These countries are not part of the preferred subset in Corak (2006) because only few estimates of the IGE are available for them, which does not allow to assess their precision. Nonetheless these 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 their IGE estimates are considerably less accurate. All the different IGE samples are reported in Table 1.

Findings of cross-country analysis. In what follows we present the results of least-square regressions of each country's IGE on various proxies of skill-substitutability. We set out 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 sample 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 the lower end, and higher end, estimates for the same core sample. The last column reports results from an expanded sample in which we add 5 extra countries to the analysis.

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-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 6 and 9 basis points (depending on the way one measures skill substitutability). These changes are not small, considering that most estimated IGE have a size between 15 and 50 basis points (see Table 1)[5]. 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 Figure 1 and 2 the scatter plots of IGE and production skill substitutability for the three different 'core' IGE samples used in the regressions reported in the first columns of Table 2. In both figures we superimpose a linear fit. The plots in Figure 1 and 2 are based on different measures of skill substitutability. As we change both the IGE samples and the way substitutability is measured, the basic patterns are remarkably similar. Countries consistently line up around the downward-sloping fit line. Crucially, countries with relatively stronger skill substitutability in production consistently exhibit higher IGE.

When we extend the core sample by including five countries (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 the estimated slope coefficients of the univariate regression using this extended sample of IGE and different measures of country-specific skill substitutability. Both magnitude and significance levels are consistent with the previous results and the scatter plots, in the right coulmns of Figures 1 and 2, confirm

[5]The R^2 values of these univariate regressions are relatively large, in some instances exceeding 40%.

a robust correlation between measures of skill substitutability and IGE levels.

3.2 Variation across US Commuting Zones

The degree of income mobility varies substantially across US regions. Chetty, Hendren, Kline, and Saez (2014) document these discrepancies and estimate different measures of mobility at the commuting zone level. We use their measures of mobility[6] to approximate the regional variation in income persistence across US localities. The disaggregation at the level of commuting zones is useful for our analysis because each commuting zone identifies a cluster of counties sharing strong commuting ties, suggesting the presence of a well-defined labor market (see Tolbert and Killian, 1987; or David Dorn, 2009).

The procedure used to generate proxies for skill complementarity for a large set of commuting zones, and to link them to mobility measures, is described in Appendix A.2. Using these data, we again find that income persistence is statistically associated with proxies of skill substitutability in production. The measures of skill substitutability at the at the level of each commuting zone are obtained by computing a weighted average of industry-specific proxies of skill substitutability, with the weights being the industry shares within each locality.

Measure of income persistence at the local level. Chetty, Hendren, Kline, and Saez (2014) estimate three basic measures to approximate the degree of mobility at the level of local "commuting zones": (1) the slope from an OLS regression of a child's rank on parental rank in the income distribution; (2) the mean percentile rank of a child whose parents are at the 25th percentile of the national income distribution; (3) the share of children reaching the top quintile of the national income distribution in their birth cohort conditional on having parents in the bottom quintile of the parents' national income distribution.

Findings of regional analysis. Table 3 reports the results of regressions of the different earnings persistence measures on proxies of skill substitutability. Each observation corresponds to an individual commuting zone. In the first column of the Table we report results for the first measure of persistence (the rank slope) capturing the difference between the expected ranks of children born to parents at the top and bottom of the income distribution. This slope measure in our sample of commuting zones takes values in the range between 2 and 50, with an average just above 30. This suggests a fairly strong rank correlation accompanied by much geographic variation. Our regressions, using the responsibility and independence O*Net proxies, indicate that a one-standard-deviation decrease in the skill substitutability measures

[6] Available at http://www.equality-of-opportunity.org/.

imply a reduction in the slope measure between .33 and 1.60: these numbers are sizeable, and correspond to drops between 1% and 5% in the average rank correlation.[7] The other two measures in Chetty, Hendren, Kline, and Saez (2014) focus on the mobility of children from families which are at the lower end of the national income distributions. In this sense they are slightly different from the first measure because they specifically focus on upward mobility. The second measure of income mobility captures the average position of a child coming from a family in the lowest quarter of the income distribution. This measure tends to assume values below 50 in our sample, with an average value around 40, suggesting that kids from families at the lower end of the income distribution tend to improve their average position. We consider this measure of income persistence in the second column of Table 3 and find that a one-standard-deviation decrease in the substitutability proxy induces an increase between 1% and 4% in the degree of upward mobility, relative to the average. Also in this case the effects are estimated with a fairly high precision. Finally, the third measure of income persistence corresponds to the probability that a child will end up in the top quintile of his cohort's income distribution, conditional on his family being in the bottom quintile. In our sample this measure has an average value of around 8, indicating that being born in a family at the bottom of the income distribution results in a fairly low probability of ending up in the top quintile of the income distribution. The third column of Table 3 shows that a one-standard deviation decrease in the degree of skill substitutability is associated to an average increase of between 2% and 10% (relative to the mean) in the probability of jumping at the high end of the income distribution.

These are very sizable statistical associations. The results confirm the findings of the cross-country analysis suggesting a co-movement between regional differences in income persistence and proxies of skill substitutability in production in local labor markets: increasing levels of skill substitutability correlates with more persistence of income. Similar correlations appear to hold for alternative measures of skill complementarity that we do not report. These results are specially interesting because we would expect them to be weaker, given that mobility of workers across commuting zones is more intense and many government policies are common across commuting zones. The proxy which can be considered closer to a general persistence measure like the IGE used in the cross-country analysis is the rank slope: in Figure 3 we plot the scatter of the rank slope measure of persistence in different commuting zones versus the two proxies of substitutability reported in Table 3 (higher proxy means lower skill substitutability).

[7] We include all 712 available commuting zones in this analysis, despite some of them clearly being outliers. Trimming the outlying observations would give even stronger results.

4 Structural Analysis

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. Both the aggregate skill distribution and the distribution 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 cross-industry allocations and the skill distribution are both stationary. Differences in primitive structural parameters across countries, which we estimate, determine differences in these steady state objects and, as a consequence, differences in intergenerational mobility.

A key feature of the model economy is its multi-sector structure, which naturally maps into the industry level OECD data. The idea is that the substitutability of skills varies across these industries, and the extent of this variation can be assessed using O*Net data. The counterfactual experiments involve re-weighing the importance of these industries in accordance with the industrial composition of alternative countries, while holding other exogenous features of the economy fixed at U.S. levels. This results in a different level of overall skill substitutability, and allows us to quantify the associated equilibrium changes in intergenerational mobility. As in the simple version of the model, the production function implies strategic complementarity in human capital investments, which creates a role for public policy that depends on the degree of skill substitutability.

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:

$$y_n = k_n^{\alpha_n} \ell_n^{1-\alpha_n}. \tag{12}$$

The capital share is allowed to vary by industry, as observed in the 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

$$\ell_n = \left(\int_{I_n} z(i) h(i)^{\lambda_n} di \right)^{\frac{1}{\lambda_n}}, \tag{13}$$

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 of 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[8]. One interpretation of the productivity shocks is lifetime labor supply shock, which do not affect the skill attainment, rather total life cycle labor supply.

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

$$Y = \prod_{n=1}^{N} y_n^{\gamma_n}. \tag{14}$$

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. We interpret the weights γ_n as arising from a combination of resource endowments, historical occurrences, geography, climate, and similar long-term characteristics, which

[8]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.

together affect relative advantage, industrial composition and aggregate technology.

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 next, 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 12 can be re-written as

$$\ell_n = \left(\int_H \int_Z z dF(z) q_n(h) h^{\lambda_n} dh \right)^{\frac{1}{\lambda_n}}, \tag{15}$$

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 weigh the measure of workers by their average productivity. Viewing the aggregate production technology as operated by a competitive representative firm, the appropriate profit maximization problem is:

$$\begin{aligned} \max_{\{q_n(h)\}} \quad & Y(\{q_n(h)\}, \{k_n\}) - \textstyle\sum_{n=1}^N \int_H \int_Z z w(h,n) q_n(h) dF(z) dh - \sum_{n=1}^N (r+\delta) k_n \\ & s.t. \\ & \{q_n(h)\} \geq \ \mathbf{0}, \\ & \{k_n\} \geq \ \mathbf{0}, \end{aligned} \tag{16}$$

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 all industry and attainment specific human capital inputs, 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 input measures:

$$q_n(h)\left[\frac{\partial Y(\{q_n(h)\},\{k_n\})}{\partial q_n(h)}-w(h,n)\right]=0. \tag{17}$$

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:

$$w(h,n)=Y\gamma_n(1-\alpha_n)\frac{1}{\lambda_n}\left\{\frac{h^{\lambda_n}}{(\int_H q_n(h)h^{\lambda_n}dh)}\right\}, \tag{18}$$

where the right-hand-side is the marginal product from equation 17. This labor demand holds whenever $q_n(h)>0$. Firms also hire capital services optimally, thus sectoral inputs satisfy:

$$\frac{\partial Y(\{q_n(h)\},\{k_n\})}{\partial k_n}-(1-\tau_k)r-\delta=0, \tag{19}$$

where δ is the capital depreciation rate. The real interest rate r is assumed exogenous, 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,

$$\ln z\sim N(\mu_z,\sigma_z^2). \tag{20}$$

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

A child's achievement, h', depends on their endowment of heritable traits $\theta'\in\Theta$, parental investments m and public investments s, through a skill formation technology $g(\theta',m+s)$. 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, 2013). We assume the same technology as that employed in Restuccia and Urrutia (2004) and earlier by Becker (1981):

$$h'=\theta'(m+s)^{\psi}. \tag{21}$$

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

$$\begin{aligned} \ln(\theta') &= \rho \ln(\theta) + \eta \\ \eta &\sim N(0, \sigma_\eta^2). \end{aligned} \tag{22}$$

Note that 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 progressivity 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 parents' decision problem can be represented recursively as:

$$\begin{aligned} V(a, h, \theta', z) &= \max_{c,m} \left\{ \tfrac{c^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}\left[V(a', h', \theta'', z') | \theta'\right] \right\} \\ s.t. & \\ c + m + a' &= zW(h)(1-\tau) + T + a(1+r) \\ h' &= g(\theta', m + s) \\ \ln(\theta'') &\sim N(\rho \ln(\theta'), \sigma_\eta^2) \\ \ln(z') &\sim N(\mu_z, \sigma_z^2). \\ a' &\geq 0 \\ W(h) &= \max_n \{w(h, n)\} \end{aligned} \tag{23}$$

Each parent has full information about her child's inherited traits, but do 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,

$$G = \tau \int w(h(i))z(i)di + \tau_k \sum_{n=1}^{N} k_n - T - s. \tag{24}$$

4.2 Equilibrium

We use the notion of stationary competitive equilibrium 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

1. The decision rules solve the household optimization problem 8, and $V(a,h,\theta',z)$ is the associated value function.

2. The representative firm optimally hires human and physical capital, thus equations 17 and 19 hold.

3. Each skill and industry specific labor market clears

$$\begin{aligned} \textstyle\sum_{n=1}^{N} q_n(h) &= \int_{A\times H\times\Theta\times Z} 1_h d\mu \ \forall \ h \in H, \\ \text{and } 0 &= q_n(h)\left[w(h,n) - W(h)\right] \end{aligned} \tag{25}$$

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

4. The goods market clears:

$$Y = \int_{A\times H\times\Theta\times Z} c(a,h,\theta',z)d\mu + \int_{A\times H\times\Theta\times Z} m(a,h,\theta',z)d\mu + G. \tag{26}$$

5. The government budget constraint in equation 24 holds.

6. 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 Equilibrium Worker-Industry Sorting

The worker-industry sorting that occurs in our model equilibrium is interesting and intuitive. Sorting is positively assortative between worker skill and industry 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. By sorting into these industries the lower skilled workers are pushed industries with greater skill complementarity. We capture this in a proposition that says that for any two workers with differing 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 that 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 Model Parameterization

Preferences. We set the risk aversion parameter $\sigma = 2$, and the discount factor to $\beta = 0.5$. This discount factor reflects the time gap between a child's outcomes and those of their parent. Based on a 25 year gap, the annualized discount factor implied by our parametrization is 0.972.

Government. 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 8). As 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:

$$\frac{Var\left[\ln\left((1-\tau)zW(h)+T\right)\right]}{Var\left[\ln(zW(h))\right]}=0.61. \tag{27}$$

We calibrate s to match the 5.5% share of GDP spent publicly on education in the US.

Industry-specific physical capital. The quantity of physical capital in each industry depends on the capital share of that industry, α_n, and on the (exogenous) gross return on capital $r+\delta$. We set the depreciation and real interest rates so that the annualized rates are 6.0% and 3.5%, respectively, and the tax rate on capital income to $\tau_k = 0.4$. The 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 value of output from each industry is equal to its aggregation weight γ_n. Thus, the aggregation weights can be parameterized by setting them equal to the share of total output attributed to each industry. We average shares observed in STAN OECD data across years 2001 to 2005.

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 can be achieved precisely by setting 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.

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

$$h' = \theta'(m+s)^{\psi}. \tag{28}$$

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

Transmission of heritable traits. The degree of persistence in heritable traits, ρ, influences the degree of intergenerational income mobility observed in equilibrium. If heritable traits are highly persistent, then parents and children tend to have similar productivity in human capital production. We match the U.S. intergenerational persistence on earnings in order to identify the persistence of heritable traits. Because we work in a stationary environment we match the intergenerational correlation of earnings, which Jantti, Bratsberg, Røed, Raaum, Naylor, Osterbacka, Bjorklund, and Eriksson (2006) estimate to be 0.357. In our quantitative experiments we scale intergenerational correlations by 1.32 to convert them to elasticities.

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

$$M_T = \frac{k - tr(P_k)}{k-1}. \tag{29}$$

This measure, which is estimated to be 0.86 for the 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 $M_T = 0.89$.

Industry-specific elasticity of substitution. We use our O*Net indicators in the parameterization of complementarity by specifying the linear relationship $\lambda_n = a_1 + a_2 \cdot$ O*Net. Our identification strategy relies on model implications for the relationships between a_1, a_2 and earnings variation: the degree of cross-sectional variation and the relationship between intra-industry variation and our O*Net indicators. To understand how the matching these moments identifies a_1 and a_2 consider the following expression for the variance of log earnings within an industry n:

$$Var_n = Var(\ln(z)) + (a_1 + a_2 \cdot \text{O*Net}_n)^2 Var_n(\ln(h)), \tag{30}$$

This can be derived from the labor demand (wage) equation with λ_n replaced by the linear specification. The variance of earnings in every industry will rise if a_1 increases, thus a_1 can be identified by matching overall earnings dispersion in the model to the data. Crucially, the log earnings variation to be matched is that for lifetime earnings, which Bowlus and Robin (2012) suggest is 30% lower than the 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_2 we replicate the relationship between within-industry wage dispersion and industry O*Net scores implied by equation 30. In CPS data we compute the standard deviation of log earnings within each industry, and then regress these on the industry level O*Net scores. The slope coefficient in this regression is 0.56. In the spirit of indirect inference, we repeat this regression using model simulated earnings data and specify a_2 so that the slope coefficient is replicated.

Summary. Table 4 presents a summary of the parameterization, other than industry level parameters. Table 5 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, while among the high substitutability sectors are education, health care and finance. These observations are consistent with the model implication that highly-skilled workers would sort to the industries in which skills are the most substitutable, and hence the returns to human capital are the highest[9].

4.5 Properties of the Benchmark Model

Sources of Persistence. In this model there are endogenous and exogenous drivers of intergenerational income persistence. The exogenous source is the persistence of heritable traits across generations, and the endogenous source is the persistence of human capital attainment across generations. To decompose these we experiment with eliminating the persistence of heritable traits, holding their variance constant. In equilibrium the IGE is reduced to 0.324, thus about 1/3 of intergenerational persistence is due to the exogenous transmission of traits, and 2/3 is due to the endogenous persistence of human capital investments.

Assessing the Effects of Education Spending. To provide some external validation of the model we assess its behavior relative to the findings of Restuccia and Urrutia (2004). In particular, we replicate their experiments on increased public education expenditure. Their model distinguishes between college and lower education. We perform identical 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. They estimate 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 expected IGE reduction from 0.4 to 0.378 in response to a 20% increase in education spending. For comparison we

[9] We have checked how average years of education by industry, measured using CPS data, relate to these estimates. Indeed, we also find a strong positive relation between educational attainment and skill substitutability in those data.

estimate the effect of a 20% increase in public education spending in our model, paid for with higher labor income taxes: this would result in a reduction of the IGE from 0.47 to 0.443. The proportional reduction in our experiment is roughly 6.1%, and very similar to the 5.8% proportional reduction suggested 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 aspects of industrial composition that are observed across countries and can be implemented in the experiments. The first is to change the relative importance of an industry in the overall economy, which is reflected in γ_n, the share of final output paid to that industry. This aspect is important because if weight is shifted to a particular industry, overall skill substitutability will rise or fall depending on whether skills in that industry are relatively substitutable or complementary. The second aspect relates to the capital intensity in different industries. If skills are quite substitutable in 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 relatively influential.

In addition to experiments that adjust output weights and capital shares, we also do experiments in which observed differences in tax and education policy are adopted. These serve two purposes: first, they provide a benchmark for comparison of the magnitudes of the effects of differences in skill complementarity; and second, they provides an upper bound on the indirect effect of skill complementarity differences (i.e. the effect of differences in incentives for public education spending).

All of our experiments are equilibrium experiments, in which a new set of wages and other equilibrium objects is attained after the U.S. equilibrium is perturbed by industrial and/or policy changes. One crucial constraint relates to the government budget identity. As overall skill substitutability changes incentives to invest in human capital, the equilibrium skill distribution will also change. This will alter government tax revenue, resulting in an unbalanced government budget constraint unless offsetting policy changes take place. We do not change the marginal tax rate as this would clearly alter the return to human capital; 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 27 is satisfied.

Table 6 provides the results of our experiments by country. For each experiment the results are divided into two parts for our core and core+5 samples. For each of these samples we provide the ratio of the standard deviations of predicted IGEs and observed IGEs, as well as the correlation between observed and predicted IGEs. The ratio of standard deviations indicates how much of the observed variation is explained by the mechanism(s) allowed for in that experiment, and the correlations are intended to show how well predicted deviations from US mobility align with actual deviations.

Focus first on the experiment in which output shares are adjusted to reflect national data, holding all other exogenous features of the economy constant. There are some great successes in explaining observed IGEs and also some failures. The experiment explains more than half of the difference in IGEs between Japan and the US, but at the same time explains very 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 shares differences account for 16% of IGE variation in the core sample and 18.5% in the core+5 sample.

When capital shares are adjusted as well as output shares the results are generally enhanced. For example, this experiment explains twice as much of the difference between the US and Denmark as when only output shares were adjusted, although 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, first compare the results of the last experiment, where only observed policy differences are applied, to the results of the first two experiment previously discussed. In the core sample observed differences in public policy explain nearly an identical amount of the variation in IGEs, as shown by the relative standard deviations. However, the correlation between predictions based on public policy and actual IGEs is one-third the size of the correlations based on the first two experiments (based on complementarity differences). Thus, if we measured effects as the product of the correlation and relative standard deviations (a regression coefficient) we would find that the effect of differences in industrial composition is three times the size of the effect of differences in public policy. In the extended sample the ratio of standard deviations is relatively large at 0.26, but this is entirely generated by counterfactually large predicted IGEs for Korea, Japan and Australia.

In the third experiment we change output shares, capital shares and public policies all at the same time to gauge the combined effects. This can be viewed as an upper bound on the total effect of differences in skill complementarity on mobility that accounts for the 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. We 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 allowing for indirect effects up to 22%.

6 Conclusion

Large discrepancies exist in the degree of social mobility across countries, yet little is known about the origins of such differences. This paper explores the extent to which alternative production arrangements may help rationalize differences in intergenerational earnings mobility across countries. Using a simple two-generation model we illustrate how the intensity of strategic complementarities in skill investments may lead to more or less earnings' mobility. We show that in this simple setting stronger skill substitutability directly dictates a higher degree of intergenerational persistence. Moreover, this analytic example suggests that a stronger skill substitutability in production should also be associated to less progressive public policies which equalize skills in the workforce.

To provide some evidence in support of this hypothesis we study geographical variation in income persistence. First, we present a descriptive analysis of international data, linking skill substitutability in production and intergenerational mobility at the country level. We find a strong and significant cross-country association between different measures of skill substitutability and estimates of IGEs. Second, we look at the geography of income persistence within the US, using estimates provided by Chetty, Hendren, Kline, and Saez (2014), finding a significant relationship between skill substitutability and income persistence at the level of 'commuting zones'. Perhaps unsurprisingly the relationship appears to be stronger at the country level, as worker can more easily relocate across commuting zones and this attenuates the effects of skill substitutability at the a finer level of local labor markets.

To explore the origins of these statistical associations we develop and estimate a structural equilibrium model. This allows us to broadly quantify the importance of skill substitutability for cross-country differences in intergenerational mobility. In the model we explicitly allow 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 perform various counterfactual experiments, which involve re-weighing 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 assess the relationship between public education policies and skill complementarities. 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 suggests the presence of an indirect effect through which technology may affect intergenerational mobility, namely by affecting incentives to adopt policies which equalize skills in the working population. When accounting for these indirect effects the model explains roughly 1/3 of cross-country differences in earnings' mobility, and can rationalize the observation of significant and persistent geographic differences in the degree of mobility and in the progressiveness of policies.

References

ABBOTT, B., G. GALLIPOLI, C. MEGHIR, AND G. VIOLANTE (2013): "Education policy and intergenerational transfers in equilibrium," Discussion paper, National Bureau of Economic Research.

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

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.

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," Discussion paper, National Bureau of Economic Research.

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.

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 (2013): "Household Choices and Child Development," *Review of Economic Studies*, Forthcoming, –.

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*.

KLEIN, P., AND G. VENTURA (2009): "Productivity differences and the dynamic effects of labor movements," *Journal of Monetary Economics*, 56(8), 1059–1073.

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

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.

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

In this paper, we use data provided by the OECD Statistics to approximate for the relative importance of different industries in a set of countries. The data used is from the structural analysis STAN databases for value added and total employment. The analysis consists of data from 31 categories of the industries, calculated as percentages of the total. The following table shows the specific industries that we have used.

A.2 Mobility and Industry Composition across the US: Data and Measurement

We use publicly available County Business Pattern (CBP) data which provide employment counts by county by industry. In particular, we use David Dorn's procedures to clean the data files and generate consistent head-counts for each commuting zone and industry pair./footnoteFor a reference, see David Autor, David Dorn and Gordon Hanson. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." American Economic Review, 103(6), 2121-2168, 2013.

The next step is to link the ISIC codes corresponding to the corresponding NAICS industry categories using the CBP data. The NAICS codes in the Dorn's files adopt the 1997 classification codes, while the ISIC codes are from 2002. We only have a crosswalk from 2002 NAICS codes to 2002 ISIC labels, therefore we need to perform an intermediate step to link 1997 NAICS to 2002 ISIC. More specifically, the Census Bureau provides a crosswalk from 1997 NAICS codes to 2002 NAICS codes,[10] as well as a crosswalk between 2002 NAICS codes and 2002 ISIC codes. Using these files we construct a correspondence between 2-digit ISIC codes and 1997 NAICS.

For each ISIC industry we generate skill complementarity measures. We first merge the ONET complementarity proxies with the occupations in the 5% Census sample for the year 2000. Then, we identify the median of each complementarity measure from the distribution of occupations within each ISIC category and we rank ISIC industries in terms of these median complementarity measures.

The resulting file is linked to the CBP data file, which gives measures of employment in

[10]We make sure that the correspondence between these codes is one-to-one. Details of this industry mapping are available.

different industries for each commuting zone level. That is, for each commuting zone we have the number of employees in industries classified into ISIC categories. It is then possible to link, for each commuting zone, the skill complementary measures (both median scores and relative ranks) with the CBP data. The 'local' skill complementarity proxies are obtained by computing a weighted average of the industries' complementarity ranks, using as weights the number of workers employed in each industry within a commuting zone.

Finally, we merge the mobility measures, provided by Chetty, Hendren, Kline, and Saez (2014), with the data on commuting zones. In this way we have a data set containing measures of mobility and skill complementarity for a large set (712) of US commuting zones. More specifically, for each commuting zone observation we have several mobility measures, NAICS and ISIC industry codes with corresponding employment counts and various proxies of skill complementarity based on O*NET measures.

A.3 Taxes and Public Education Spending

Table 8 reports public education expenditure levels and labor income tax burdens, primarily based on OECD data sources.

A.4 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 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 (30), 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[11] 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 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

[11] 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.

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 4 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 9. In this Table we replicate the exercise illustrated in figure 4 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.[12] Table 9 reports the resulting slope coefficients and standard deviations for OLS regressions with robust standard errors.

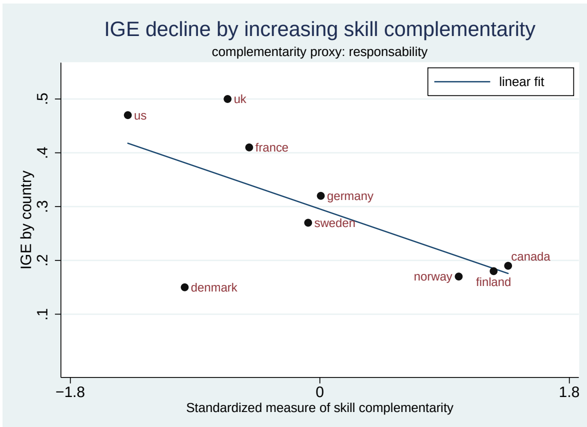
The results in Table 9 show that the industry ranks calculated for all countries in the sample are highly correlated. As in the case of Germany, the correlation for ranks of labor income dispersion is stronger than for wage dispersion. Yet, in the whole sample, controlling for demographic characteristics 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

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

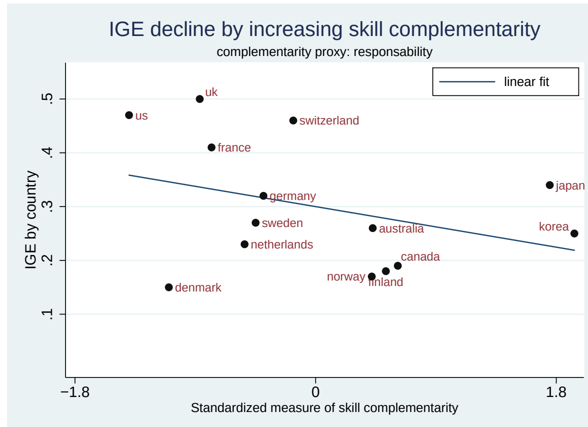
pattern of within-industry wage dispersion must be due to some aspect of industry structure.

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 technology 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.

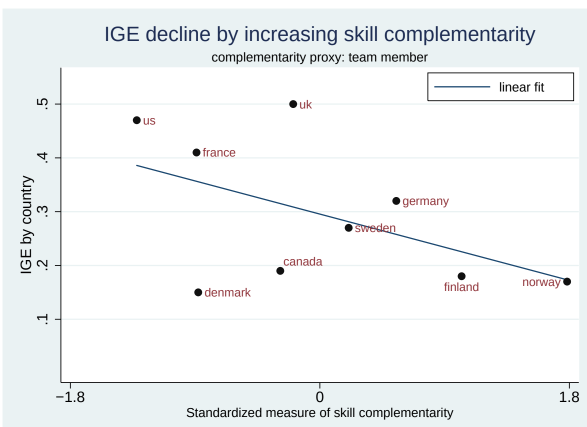
Figure 1: IGE vs average skill substitutability in the core IGE sample (left column) and in the extended sample (right colum). ONET proxies: (top row) responsibility; (middle row) team member; (bottom row) common factor of four separate ONET proxies.



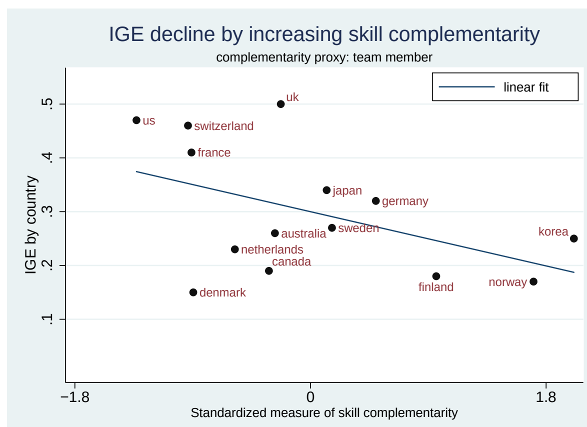
(a) core IGE sample, responsibility proxy



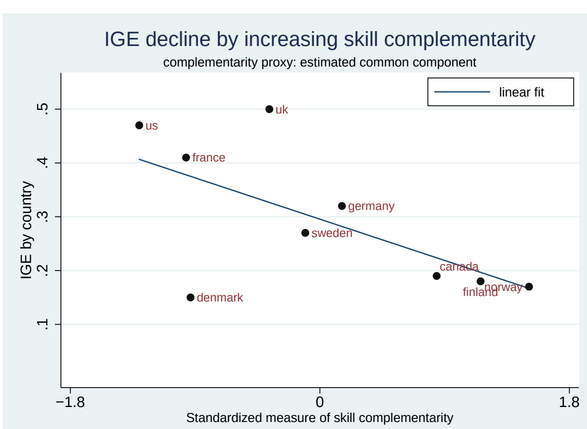
(b) extended IGE sample, responsibility proxy



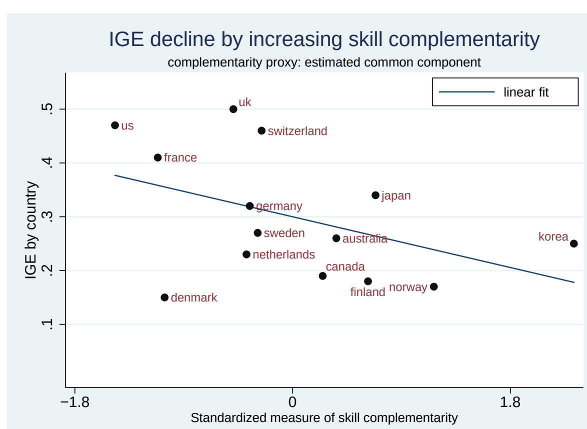
(c) core IGE sample, team member proxy



(d) extended IGE sample, team member proxy

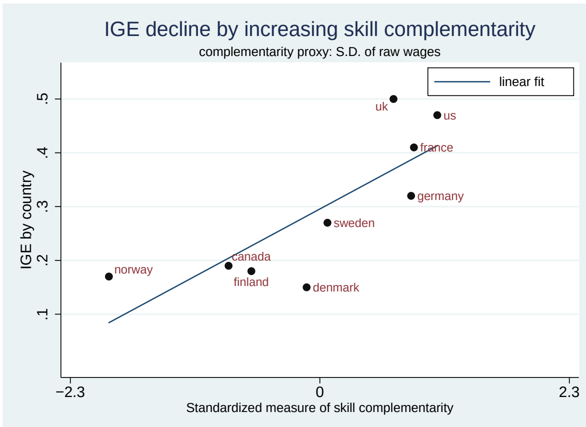


(e) core IGE sample, common component proxy

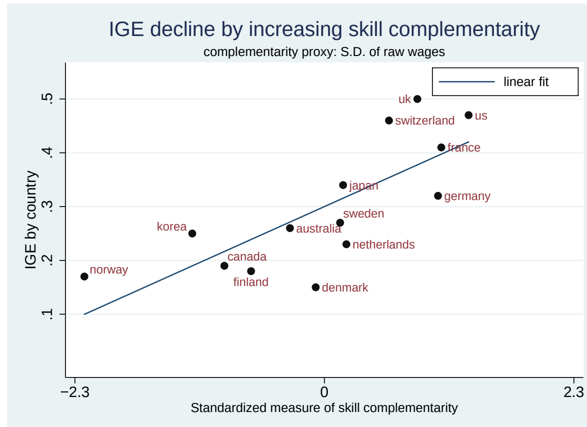


(f) extended IGE sample, common component proxy

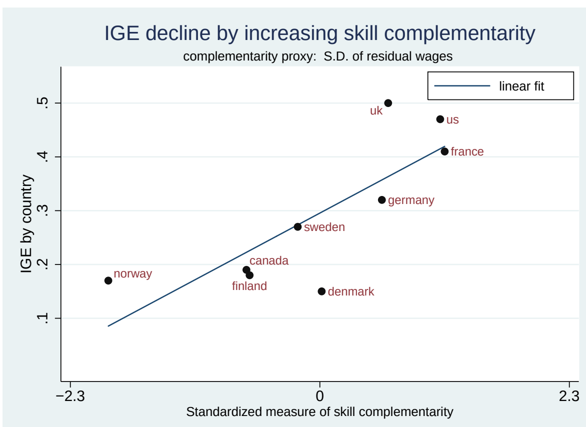
Figure 2: IGE vs average skill substitutability in the core IGE sample (left column) and in the extended sample (right colum). Wage dispersion proxies: (top row) S.D. of raw wages; (middle row) S.D. of residual wages; (bottom row) common factor of four separate wage dispersion proxies.



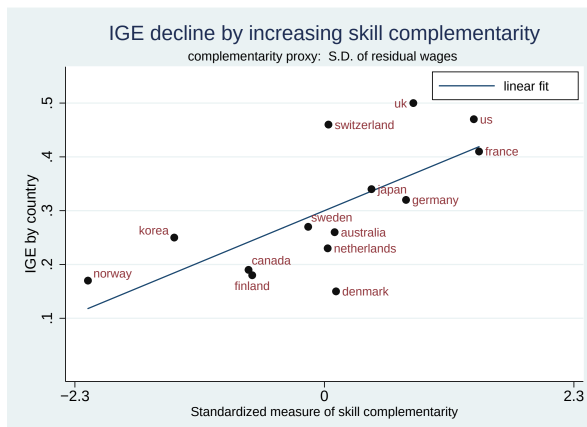
(a) core IGE sample, S.D. of raw wages proxy



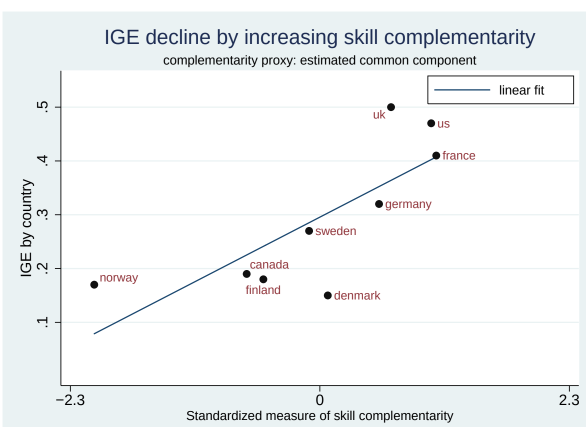
(b) extended IGE sample, S.D. of raw wages proxy



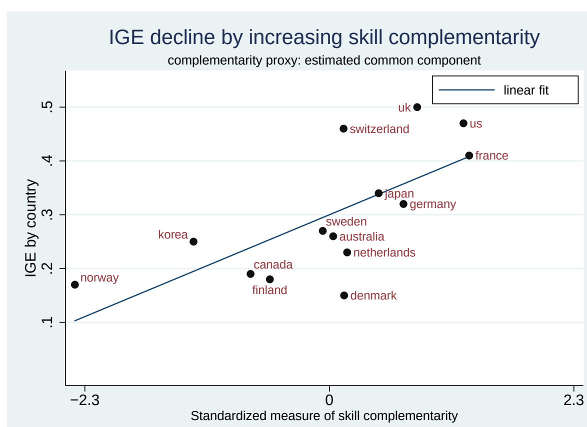
(c) core IGE sample, S.D. of residual wages proxy



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

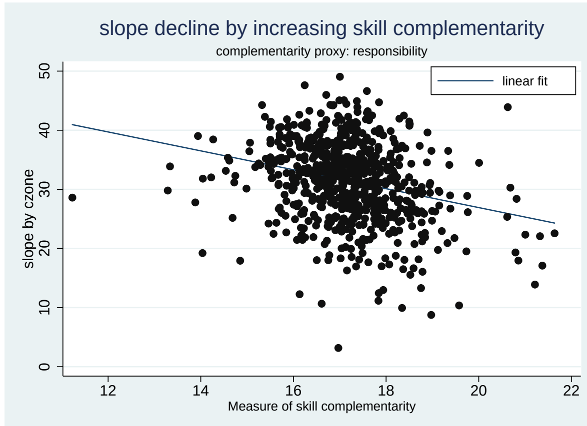


(e) core IGE sample, common component proxy

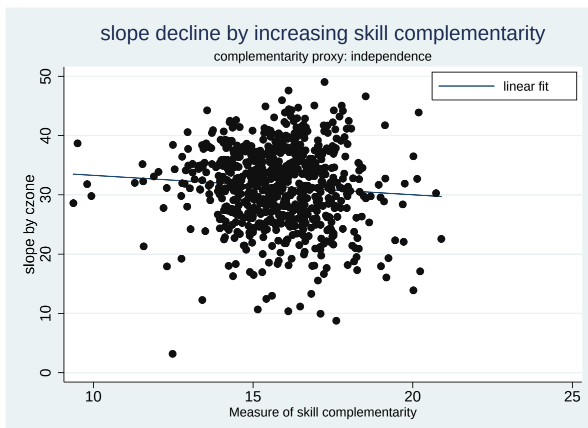


(f) extended IGE sample, common component proxy

Figure 3: Income persistence (rank slope) vs skill substitutability proxies (proxies decrease in substitutability).

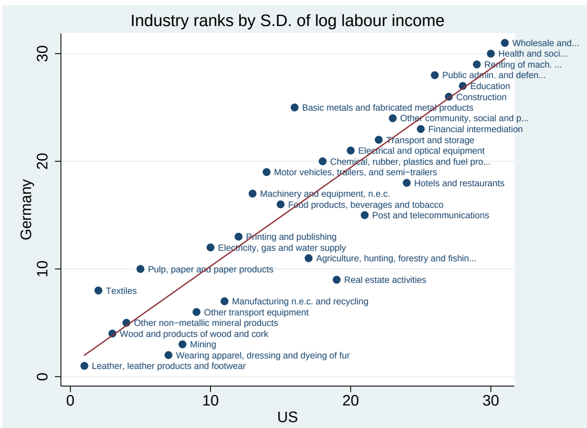


(a) persistence (rank slope) vs responsibility proxy

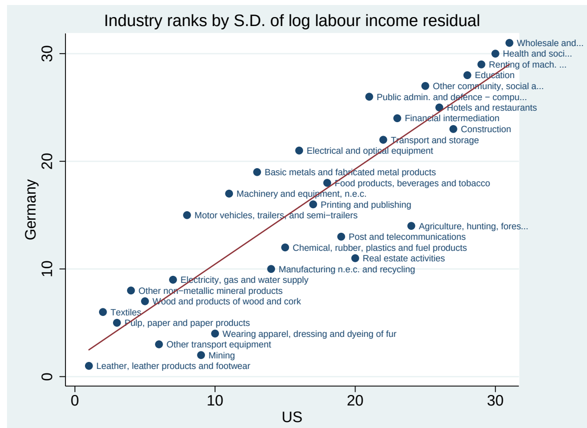


(b) persistence (rank slope) vs independence proxy

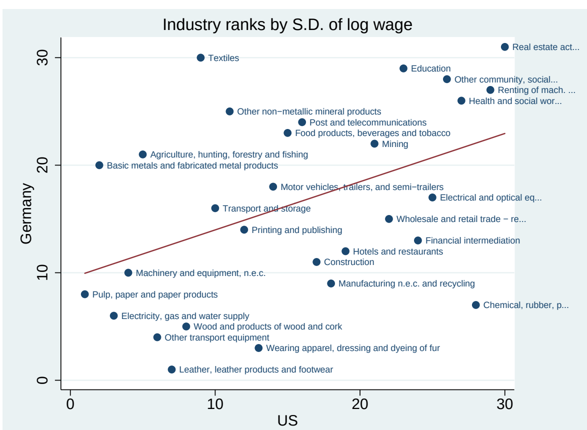
Figure 4: Comparison of US and German industry ranks according to different dispersion statistics. "Residual" refers to the remaining unexplained variation after controlling for industry, education, age, sex, region (state or province), and year.



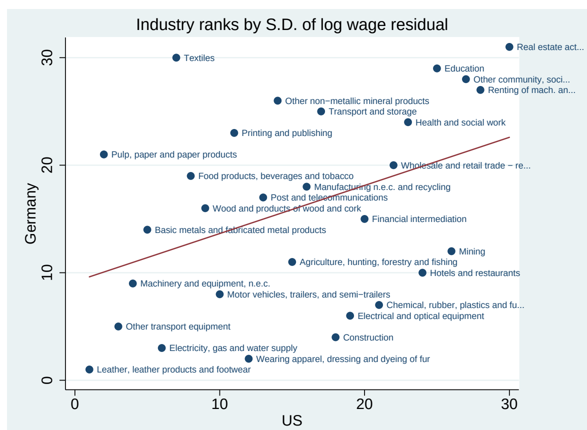
(a) Log labour income dispersion ranks



(b) Log labour income residual dispersion ranks



(c) Log wage dispersion ranks



(d) Log wage residual dispersion ranks

Table 1: Different samples of IGE estimates.

	core	core (low end)	core (high end)	core + 6
	(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

Table 2: Estimated change in the level of IGE associated to a (one standard deviation) decrease in the skill-substitutability proxy. The first panel uses proxies based on ONET measures; the second panel uses proxies based on wage dispersion. Each column corresponds to a different sample of IGE estimates (the first three samples include only the nine core countries; the last one adds five more countries to the core sample). Standard errors in parenthesis.

	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	-.07	-.06	-.07	-.06
	(.03)	(.03)	(.04)	(.02)
responsibility	-.09	-.07	-.09	-.04
	(.04)	(.03)	(.04)	(.03)
independence	-.08	-.07	-.09	-.05
	(.03)	(.02)	(.03)	(.03)
estimated common component	-.08	-.07	-.09	-.05
	(.03)	(.03)	(.03)	(.03)

	IGE measures			
	core	core (low end)	core (high end)	core + 5
skill substitutability proxy	(1)	(2)	(3)	(4)
Wage dispersion proxies (increasing in substitutability)				
C.V. of raw wages	.10	.08	.11	.08
	(.03)	(.03)	(.04)	(.02)
S.D. of raw wages	.11	.09	.12	.09
	(.03)	(.03)	(.03)	(.02)
C.V. of residual wages	.09	.08	.10	.07
	(.03)	(.03)	(.03)	(.02)
S.D. of residual wages	.11	.09	.12	.08
	(.03)	(.02)	(.03)	(.02)
estimated common component	.10	.09	.11	.08
	(.03)	(.03)	(.03)	(.02)

Notes

· The common component proxies are estimated using the proxies of skill substitutability listed in each panel.

· C.V.=coefficient of variation; S.D.=standard deviation.

Table 3: Estimated change in earnings persistence associated to a (one standard deviation) increase in the skill-substitutability proxy. Each column corresponds to a different measure of intergenerational persistence: (1) is the rank regression slope; (2) is the conditional mean percentile rank; (3) is the conditional probability of being in the top quintile. Standard errors in parenthesis.

	Intergenerational income persistence measures		
	rank slope	mean percentile	bottom-to-top quintile
skill substitutability proxy	(1)	(2)	(3)
ONET proxies (decreasing in substitutability)			
responsibility	-1.60	1.40	1.18
	(.24)	(.18)	(.16)
independence	-.33	.45	.28
	(.17)	(.14)	(.11)

Table 4: This table reports the parameterization results for non-industry level parameters.

Parameter	Notation	Value
Calibrated		
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
Substitution Parameter Constant	a_1	0.504
Substitution Parameter Slope	a_2	1.801
Fixed		
Intergenerational Discount Factor	β	0.5
CRRA Parameter	σ	2.0
Net Annualized Interest Rate	r	0.03
Annualized Depreciation Rate	δ	0.06

Table 5: This table reports the parameterization results for industry specific production parameters.

Industry	γ_n Industry Share	α_n Capital Share	λ_n 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

Table 6: Our four experiments involve: (1)[Output Shares] adjusting output shares, γ_n, to reflect national data; (2)[Out. and Labor Shares] adjusting both output and capital shares, γ_n and α_n, to reflect national data; (3)[Obs. Policies and Shares] adjusting output and capital shares as well as education subsidization rates s and labor income tax rates τ to reflect national data; and (4)[Obs. Policies only] adjusting only public education and tax rates to reflect observed national data. Note that capital share data is not available for Switzerland, thus US capital shares are maintained.

Country	Literature IGE Estimates	Output Shares	Out. and capital Shares	Obs. Policies and Shares	Obs. Policies only
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	–	0.668	0.509	0.596	0.204
Relative s.d.	–	0.160	0.171	0.174	0.160
Core + 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	0.46	0.441	–	0.456	0.481
Correlation	–	0.430	0.428	0.631	0.285
Relative s.d.	–	0.185	0.205	0.220	0.260

Table 7: This table reports the industries used to decompose employment at the country level. 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
âĂć	Residual (all remaining industries)

Table 8: The second column of this table reports the percentage of GDP that is spent by all levels of government on education, as per the OECD document "Education at a Glance 2013." *German data taken from World Bank data. The third column reports the percentage of labor earnings that is paid as income/payroll taxes or social security contributions as reported by the OECD.

Country	Public Education Spending as % of GDP	Taxes as % of Labor Income
Core Sample		
United States	5.5%	29.6%
Canada	5.3%	30.8%
Denmark	8.8%	38.6%
Finland	6.8%	42.5%
France	5.9%	50.2%
Norway	8.8%	37.6%
Sweden	7.0%	42.8%
Germany	5.1%*	49.7%
UK	6.3%	32.3%
Core + 5 Sample		
Australia	5.2%	27.2%
Japan	3.8%	31.2%
Korea	4.9%	21.0%
Netherlands	6.0%	38.6%
Switzerland	5.2%	21.5%

Table 9: This table reports the correspondence of within-industry income and wage dispersion patterns across countries. 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. Then the ranks of each country and dispersion statistic are regressed independently on the corresponding ranks calculated 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 wage.

	Industry S.D. rank regression slope coefficients			
	Dependant variable: industry S.D. ranks, pooled US 2004,2007,2010			
	log labour income	log labour income residual	log wage	log wage residual
	(1)	(2)	(3)	(4)
Pooled country-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 country-years				
US 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)		