Paper to be Presented at the
RC28 Spring Meeting in Brno, Czech Republic, May 24-27, 2007

A Paradoxical Latent Structure of Educational Inequality: Cognitive Ability and Family Background across Diverse Societies[1]

Tony Tam, Academia Sinica, Taiwan

ABSTRACT

The objective of this paper is to identify the cross-national structure of family influences on the cognitive achievement status of children. The study is guided by a focus on positional inequality and three theoretical perspectives. For empirical operationalization, this study parameterizes the degree of intergenerational closure in educational achievement in a path-analytic framework and applies the method of positional status index (PSI) to overcome two problems of comparability in cross-national research. The methodology results in a ratio-scale metric of family effects that is identically interpretable across four cognitive outcomes in 41 countries. Empirical analysis is based on PISA 2003 (cross-national surveys of 15-year old students, with standardized tests on problem solving skills, reading, math, and science literacy). The PSI metric enables a latent class modeling approach to cluster analysis and nonparametric factor analysis, and results in three striking findings: (1) There is an intriguing and sharp differentiation of the countries with respect to closure. The degree of closure is highly variable but the latent structure is a simple one-dimensional concept of closure in positional competition. This dimension is tightly connected with parent education effects and the combined effects of parental education and household socioeconomic status. (2) However, there is also surprising heterogeneity among societies occupying the same level of closure, with apparently very different educational histories, cultures, political institutions, and levels of socioeconomic development. For instance, post-communist states (such as Hungary) are disproportionately represented among societies with strongest family background effects. By contrast, societies with the lowest levels of family effects are capitalist societies. (3) Taken together, the cross-national latent structure of family effects is most compatible with the positional competition perspective, vindicating the usefulness of the focus on positional inequality.

Keywords: Comparative educational stratification, positional competition, cognitive achievement, family background, latent class modeling.

---

[1] This research was supported by Academia Sinica through the Learning 2000 Project and the National Science Council (NSC 94-2412- H-001-025). Direct correspondence to Tony Tam, Academia Sinica, Institute of European and American Studies, Nankang, Taipei, 11529, TAIWAN (email: tam@sinica.edu.tw).

A Paradoxical Latent Structure of Educational Inequality: Cognitive Ability and Family Background across Diverse Societies

Tony Tam

I. Research Questions

Sociology has a long and strong tradition in studying the role of cognitive achievement in the process of stratification and how it mediates family influences on a person's life chances. The golden period of quantitative analysis of cognitive inequality and family influences dates back to the 1970s when a series of landmark studies (Jencks 1972, 1979; Sewell and Hauser 1975; Hauser and Featherman 1977; Featherman and Hauser 1978) were conducted in response to the seminal contributions of Coleman (Coleman et al. 1966) and Blau and Duncan (1967). Since the mid 1990s, the Bell Curve debate spurred by Herrnstein and Murray (1994) and the apparent rise in returns to cognitive skills have led economists to join sociologists in the study of cognitive inequality (see Heckman, Krueger, and Benjamin [2003] for a comprehensive and recent review of original studies). In line with the concerns of this literature, this study examines the fundamental relationship between family origin and cognitive inequality using large-scale multi-country cross-sectional surveys that include comparable cognitive test scores.

Despite the enormous progress made, it remains difficult to synthesize country-specific results from diverse national and historical contexts. The reason is surprisingly simple: the striking absence of a statistical modeling and measurement framework in which comparative research on cognitive inequality can be conducted

without making heroic assumptions on the unknown distribution of cognitive skills in a population (e.g. cognitive ability universally follows a normal or logistic distribution), how the ability distributions may be similar or different across populations, or how different categories of parental education and income can be compared across societies with different currencies and educational system. To be sure, existing frameworks do not allow results from different contexts to be readily compared and synthesized.[2] These impediments have seriously limited the types of research questions raised and the efficiency of cumulative research progress.

The purpose of this study is to identify the structure of inequality in cognitive skills that are fundamental to social, economic, and cultural stratification across a diversity of modern societies. The punch line is that there is indeed a surprisingly simple structure of inequality to be discovered. However, the discovery of the structure requires overcoming the obstacles that impede previous comparative research. This paper therefore presents the fruit of the search for a simple structure and the solutions I developed to make the discovery possible.

The search is guided by a theoretical framework that revives a focus on positional inequality (as opposed to absolute inequality) and compares the predictions of three perspectives. To take on the methodological obstacles, I introduce two methodological tools. The first one is to operationalize the educational closure of a society with a path-analytic framework. The second one is motivated by the problems of incomparability that plague cross-national measurement. The tools allow the formulation of a succinct test of three theoretical perspectives on cross-national

[2] Hout and DiPrete (2006, p. 6-7) discuss the problem in the context of the relative absence of attention to comparative studies of Blau and Duncan's (1967) seminal model specification of the roles of education in the status attainment process.

variation in family influences. An ideal database for this study is PISA 2003, a major set of cross-national[3] student surveys to conduct a new empirical test of three theoretical perspectives.

II. Theoretical Framework

Positional Inequality

The first theoretical premise of this study is that there is much to be learnt from a sustained analysis of rank-based inequality in the search for the structure of intergenerational influences on educational achievement. Indeed, the starting point of the conceptual framework is the distinction between absolute inequality (such as the size of gender earnings gap) and rank-based inequality (such as the sign of the gender earnings gap that signals which gender gap has an earnings advantage over the other group). A society may have witnessed a large reduction in absolute income inequality without altering any existing rank-based inequality of income. Rank-based inequality is what the literature has come to call positional inequality.

Scholarly focus on positional inequality is not new. It has been central to a prosperous variety of theories of inequality and stratification advanced in the 1970s, such as the vacancy chain mobility theory of White (1970), queuing theory of Thurow (1975), job competition theories of Sorensen (1977) and Thurow (1975), positional status theories of Hirsch (1976) and Frank (1985), and the rank-order tournament model of Lazear and Rosen (1981) for internal labor markets. Educational achievement is one of the most pervasive forms of positional inequality.

The focus on positional inequality also serves a strategic methodological

[3] Strictly speaking, not all the participating members of the surveys are countries *per se*. Nevertheless,

purpose in the quest for the structure of cross-national variation in family influences on educational achievement. The purpose is to reduce the problems of incomparability that often plague cross-national research. As will be clear, the methodological solutions adopted in this study to resolve incomparability problems are closely related to a focus on positional inequality.

Three Perspectives

Where does cognitive inequality come from? After decades of controversies (such as *The Bell Curve* debate), the state of the art assessment of the cumulative evidence from a wide range of research design strongly confirms the view that, apart from a biological (in fact, genetic) component, cognitive skills are to a large extent nurtured. The extent and quality of nurturing is what sociologists often attribute to as family and schooling influences whereas economists often call it the quantity and quality of investment in human capital. Early large differentials of investment in cognitive skill formation result in large and persistent cognitive skill differentials later in life (Carneiro and Heckman 2003).

What theoretical guidance is available on the social origin of cognitive inequality? While there are fairly well elaborated theories of educational and status attainment, sociological theories of cognitive inequality are much more limited. The literature on family size effects and birth order is, as it stands, more driven by psychological and human capital theories than by sociological ones. The starting point of this study is the recognition that three of the major theoretical perspectives originally developed for labor market inequality have implications for cognitive

---

they are administratively and socially distinct societies that resemble a country.

inequality as well: (1) education as a socio-cultural institution for socioeconomic development, (2) education as a means of group domination via opportunity exclusion, and (3) education as status production via interpersonal and inter-family competition.

Each perspective encompasses a range of familiar theories and conjectures. The way I define or group the existing theories into the perspectives is informed by the substantive context of the study. A major representative of the first perspective is modernization theory and related functionalist theories of social organization and social change, especially over the 20th century (Parsons 1970; Treiman 1970). The second perspective is mostly associated with theories that emphasize social conflict and the struggle for one group to dominate other groups, and therefore the perpetuation or reproduction of inequality from the one generation to the next (Collins 1971, 1979; Bourdieu 1973). The third perspective is based on positional competition theories motivated by observations such as the importance of relative standing and people's instinctive pursuit of higher relative standing (Boudon 1974; Thurow 1975; Hirsch 1976; Sorensen 1977; Frank 1985), leading to such phenomena as overeducation and the inflation of educational credentials.

Hypotheses

Empirical research in social mobility has been for a long time largely descriptive and exploratory rather than strongly driven by theory testing (for, e.g., Breen and Jonsson (2005) for a recent review). Researchers are often flooded with rich comparative and descriptive results, yet barely able to come up with sharp theoretical explanations. This is, unfortunately, also the case with educational mobility research. This study offers a small step forward in the context of comparative

educational stratification research.

The present study is guided by three theoretical perspectives: socioeconomic development, social conflict, and positional competition. The first perspective predicts a distinctively negative correlation between total family effect and the level of socioeconomic development or modernization of a society. Under this perspective, *total family effect on educational status is viewed as a measure of the educational closure in a society, hence it should be positively related to the level of socioeconomic development or modernization of the society.*

By contrast, the second perspective expects social class to be endemic to all societies. Capitalist societies are all similar in nature and, as Erikson and Goldthorpe's (1992) influential work suggests, major industrial societies display a strong convergence in their class mobility patterns. Capitalism and communism/socialism represent two ideal types of modern class regimes. To the extent that class regimes dictate not only occupational mobility but also cognitive inequality, the perspective predicts *the tendency toward capitalist-communist bifurcation of family effects on cognitive inequality*.

The third perspective entails convergent effects of parental competitive success on the cognitive achievement of children and competitive success is a continuous rather than categorical phenomenon. What institutional differences are crucial for defining the rules of competition in educational success? Drawing on the latest theoretical and empirical advance on the formation of human capital (Carneiro and Heckman 2003), the perspective takes seriously the view that parental education rather than parental class and extra-family intervention is the crucial engine of human capital formation. The strength of family influences is bound to vary across societies,

depending on the role played by schools and other social institutions outside the family. Schooling systems differ, for instance, by the extent to which elementary and secondary schools serve to equalize or aggravate stratification of students from diverse family backgrounds. How society enters into the picture is not driven by socioeconomic development or the structure of power in politics, but by the social choice of a society with regard to how much the society is willing to intervene to equalize learning opportunities of students from different socioeconomic backgrounds. *The prediction is therefore that a simple, even one-dimensional differentiation of societies in terms of educational closure is possible and the level of closure is not determined by socioeconomic development or polity per se*.

III. Analytic Framework

Parameterizing Closure

The first conceptual tool is a way to think about the "openness" of a society for intergenerational mobility in educational achievement. The converse of this openness is intergenerational "closure"—the degree to which the educational achievement of the child generation is determined by the socioeconomic standing of the parent generation.[4] Panel A of figure 1 presents the basic conceptual framework. Panel B is a path-analytic device for conceptualizing the role of family vs. society in the generation of educational achievement status. The device is an extension of the path model that is central to understanding the conceptual and empirical contributions of

[4] While it is equivalent to think in terms of openness, the notion of closure has the advantage of a direct correspondence with the parameters of interest to be examined in this paper.

the Blau and Duncan (1967) treatise on intergenerational occupational mobility.[5] The π-parameters represent effects of socioeconomic background, the β-parameters are the effects of skill formation processes within and outside the family on the educational achievement of a person, and the ε-parameters reflect the equalizing influences of exogenous forces, such as state intervention, that disrupt the rank-order stability of intergenerational mobility. Note that the ε-parameters are closely related to the variance unexplained ($1-R^2$).

INSERT FIGURE 1 ABOUT HERE.

The simple parametric structure of the path-analytic representation offers powerful clarity in defining some important concepts for the question of intergenerational mobility in the context of cognitive achievement as much as it does in the context of intergenerational occupational mobility. Three notions are particularly crucial: (1) The ratio of $\pi_S\beta_S$ to $\pi_F\beta_F$ measures the extent to which **S** plays a reproducer role for parental socioeconomic inequality. (2) As $\varepsilon_S\beta_S$ and ε_Y would reflect the equalizing effects of non-family influences (social intervention and so forth), the ratio of $\varepsilon_S\beta_S$ to $\pi_S\beta_S$ determines the extent to which **S** plays an equalizer role vis-à-vis an reproducer role. (3) In addition, the ratio of $(\pi_F\beta_F + \pi_S\beta_S)$ to $(\varepsilon_S\beta_S + \varepsilon_Y)$ is a good indicator of the overall closure of a society for educational achievement. This framework simultaneously operationalizes the concept of reproduction/closure and offers a decomposition of its sources.

However, all large scale cross-national data to date do not have rich enough data to fully identify the model without imposing some a priori assumptions on the

[5] Note that figure 1 amounts to a decomposition of the origin-education link in the basic Blau and Duncan model, except that origin here is restricted to parental status as derived from education and income and the destination here is cognitive achievement status rather than occupational status.

model. The second and third notions are particularly difficult to identify. But we do not have to be overly pessimistic. In fact, even the neoclassical economic growth model of Robert Solow, which ushered in a revolution to the macroeconomics of growth and development economics, was first proposed as a model full of simplifying and identifying assumptions. We just have to keep in mind that identifying assumption is an analytic tool that, unlike an epistemological foundation, can be adaptively respecified and revised as the data permit and the evidence warrants.

A powerful identifying assumption for the model in figure 1 is to assume that all societies in the data share essentially the same technology of skill formation within the family. In other words, the direct effect of family influences ($\pi_F\beta_F$) is by and large constant across societies. The source of variable family influences therefore resides in the indirect effect of family influences mediated by the skill formation institutions outside the family ($\pi_S\beta_S$).[6] This implication has a significant payoff: it greatly simplifies the interpretation of the cross-national results to be presented below. The "total" effect of socioeconomic background (the sum of direct and indirect effects) can be estimated with minimal data: $\pi_F\beta_F + \pi_S\beta_S$.[7] Cross-national differences of this total family effect can be interpreted as cross-national differences in the indirect effect—the extent to which the schooling system in particular and other social institutions outside the family contribute to the reproduction of family socioeconomic inequality.

To sum, high total family effect signals high closure—strong reproduction via schooling and other social institutions. In addition, the lowest observed level of total

[6] Another assumption, implicit in the literature and for most of our discussion here, is that measurement errors are the same across the sample societies in the data. Relaxing this assumption may prove necessary, but it will open up a large number of indeterminacies and demand rich data for any serious solution to the estimation problems. With sufficiently rich data, both identifying assumptions can be tested.

family effect may be regarded as an upper bound estimate of the relatively constant direct effect of family influences on educational achievement. Thus "the total family effect of society X" minus "the total family effect of the most open society in the sample" provides a rough estimate of $\pi_S\beta_S$ for society X. Nevertheless, the elegance of the path-analytic representation of the concept of closure will be void if cross-national data cannot be measured and analyzed with a good deal of comparability. We will therefore consider the basic problems of incomparability in detail and propose solutions compatible with the theoretical focus of this study.

Problems of Incomparability

With the advent of item response theory (IRT) (Rasch 1960), psychometricians have built a rigorous approach to deduce latent ability from simple right (1) or wrong (0) binary response profile on a standardized test (Lord 1980). Well-known student surveys have routinely relied on the theory to produce cognitive achievement scores as an outcome of learning. Unfortunately, two profound shortcomings of using IRT scores as cognitive achievement have greatly impeded the progress of empirical research. Sociologists and other social scientists interested in educational inequality have taken too much of the virtues and power of the IRT scores for granted. It is high time that the research community faces up to the limitations and thinks of alternative measures of educational outcome meaningful for the study of inequality.

The first problem is scale indeterminacy. The problem is well-known even in the most basic Rasch model of modern item response theory (Rasch 1960). The simplest way to understanding the Rasch model is to think of it as a special logistic

---

[7] As it is generally the case, the parameters of both paths are positive, so $\pi_F\beta_F > 0$ and $\pi_S\beta_S > 0$.

regression in which the logit of a correct response (the log odds of a correct response to a test item *i* by a person *j*) is an additive function of person-specific ability (θ_j) and item-specific difficulty (b_i):

$$\log \frac{P_{ij}(\theta_j)}{1-P_{ij}(\theta_j)} = \sigma(\theta_j - b_i).$$

The Rasch model assumes that the probability of a correct response is nonlinearly related to the ability and item difficulty via a logistic functional form. Most significant, the scale of the ability and difficulty parameters are inherently inseparable from the parameter σ and the location (implying the mean) of the ability and difficulty parameters are also unidentifiable. The choice of any positive scale for the ability and difficulty parameters would not affect the fit of the model with data. All choices of scale are equally consistent with the data, hence the problem of scale indeterminacy. By definition of the IRT framework, then, the scale and location of IRT scores are inherently unidentifiable. The dominance of the IRT framework ensures that the scale indeterminacy problem is ubiquitous in empirical studies of cognitive inequality.

With the scale of the parameters indeterminate, any model for cognitive differences is met with the unpleasant situation that the size of any effect parameter cannot be interpreted in absolute term. Only the relative effects of explanatory variables are potentially interpretable. For instance, if the estimated effect of father's years of education on cognitive ability of children is 10 whereas that of mother's education is 15, we cannot make anything out of 10 or 15, but at most infer that mother's educational effect is 1.5 times that of father's.

The second problem concerns a multiplicity of incomparability that has long

plagued comparative educational stratification research. Consider, for instance, the strong interest in estimating the relative roles played by financial and nonfinancial family factors. The distinction is significant because financial factors are amenable to policy intervention whereas most nonfinancial factors are not (Cameron and Heckman 1998, 2001). If short-term liquidity constraint or family poverty is a crucial disadvantage, the public may alleviate the adverse impact of poverty by offering financial subsidies for schooling without interfering with any behavior and process inside the family. Economists are still far from settled about the meaningfulness of a variety of criteria of evaluating relative effects (Heckman, Krueger, and Friedman 2003), even less about the optimal way to do so.

Without getting into this controversial issue, I would limit attention to identifying the relative importance in terms of structural effects. But assessing the role of financial versus nonfinancial factors in terms of structural effects has to address the problem of comparing effects between variables measured in different units. In addition, if IRT scores are inherently under-identified, how can results from separate studies and contexts be comparable? If and when the IRT scores from two studies are based on overlapped tests, it is possible to impose further assumptions in order to place the two sets of scores on the same metric and so comparable. But the assumptions are fallible and, in most cases, tests for separate studies are necessarily incompatible.

As I will argue shortly, the problem is not really insurmountable. The key to avoiding the problem of incomparability is to convert under-identified test scores into observable ratio-scale scores of cognitive achievement. To do so, I apply the positional status index (PSI) methodology to quantify the relationship between

cognitive achievement and parental background. The methodology makes possible a parsimonious operationalization of the conceptual ideas. Parental influences on cognitive achievement can be succinctly summarized by two effect parameters; and the effect parameters share a common metric and interpretation across cognitive tests, parental variable, and societies.

Positional Status Index

The methodological key to the present study is a positional status index, labeled as the PSI methodology by Tam (2005). To fix ideas, consider the paradigmatic case where cognitive achievement is precisely measured to differentiate every student from each other. Each student occupies a unique position; the distribution of N positions coincides with the distribution of N students. In this case of perfect (or similarly extreme) differentiation, it is useful to standardize the ceiling and floor of a PSI so that the range of PSI does not vary wildly across systems of very different sizes. In practice, it is adequate enough to set the ceiling and floor of PSI at 999 and .001, respectively. This range avoids differentiating students beyond the top .1 percent and basically those below the bottom .1 percent of a sample. The PSI score for position k is computed as the following odds:

$$\mathrm{PSI_k} = \frac{P_k}{1-P_k}$$

$$= \frac{1-Q_k}{Q_k}$$

where P_k is the proportion of students below level k and Q_k is level k or above. Despite the simplicity of its functional form, the PSI score richly captures the competitive structure of the schooling hierarchy and translates it into a status score for

each position. This simple form of PSI has an intuitive interpretation that underscores the social and interdependent nature of competitive success. The index represents *the average number of competitors a person has to beat or exclude* in order to reach level *k* — *a*verage *c*ompetitors *e*xcluded may be dubbed the ACE interpretation of PSI.

The ACE interpretation of PSI is also informative about the measurement properties of the metric. It highlights the fact that the metric has a tangible and observable reference — unlike any latent variable construct such as the IRT scores psychometricians use to define cognitive ability. Unlike the IRT scores, the scale or location of the PSI metric is not arbitrary; the metric is sensitive even to linear transformations. In addition, the ACE interpretation reveals the nice property that PSI is a ratio-scale metric since zero is an absolutely meaningful origin of the metric — the absence of any competitive standing. The metric meaning is *universal across contexts and over time*. Thus the PSI transformation of IRT scores avoids the scale indeterminacy problem.

As pointed out earlier, the incomparability of effects between explanatory variables of different units of measurement used to be an insurmountable problem. In the 1960s and 1970s, social scientists were used to rely on the decomposition of variance explained to evaluate the relative importance of explanatory variables. By now quantitative researchers have recognized that this technique is ill-conceived and often misleading (Hanushek and Jackson 1977, p. 59; Lieberson 1985, p. 90-91). A related and similarly misguided technique is the use of standardized coefficients (Hanushek and Jackson 1977, p. 78-79).

But as Tam (2005) has demonstrated in the analysis of Taiwanese and American data (TEPS and NELS, respectively), the PSI framework is immune to the

traditional problems of incomparability discussed above. The key is to score parental background and cognitive achievement in terms of an interpretable metric with fixed meaning for every value on the metric, such as PSI, so that the variables can be analyzed without any problem of identification. The theoretical argument for conceptualizing cognitive achievement as competitive success and therefore converting IRT scores into PSI is equally applicable to paternal and maternal education and household socioeconomic status in the production of a student's cognitive PSI. The causal significance of parental education and socioeconomic status signify the relative standing of parents in the parental generation. Conceptualizing household socioeconomic status of a student as a measure of the parental competitive success, all parents of students of a given cohort and society can be pooled to form an ordered distribution of household socioeconomic status. The PSI of household socioeconomic status can then be computed just as the PSI of parental education. With both the dependent and independent variables in the PSI metric, there is no problem of comparability between the coefficients of household socioeconomic status and education, nor problem of comparing coefficients across grades and societies.

IV. Data and Method

Data

The analytic plan of this study involves the analysis of PISA2003, a collection of large-scale cross-sectional student surveys comparably conducted in multiple countries. For the purposes and scope of the study, PISA2003 is a good source of data, even though it is limited to students at the equivalent of the end of junior high school, those just before becoming eligible to enter the labor market for most modern

societies.

PISA2003 is particularly suitable for this study. First, the design of PISA is informed by sociological research on family influences, with special attention to measuring family background. For further research, there is also a rich set of indicators on economic, cultural, and social capital. Second, PISA is consciously designed to measure cognitive ability needed for effective functioning in the knowledge economy to date. Thus cross-national comparisons are intended to be as independent of cross-national curriculum differences as possible. Third, PISA2003 offers four measures of cognitive achievement: problem-solving ability, math literacy, science literacy, and reading literacy. This is currently the most comprehensive set of cognitive measures and for a large collection of societies, both features promote dependability and robustness of the results.

For each cognitive test, two regression models are estimated for each society. The first model examines the total effects of paternal and maternal years of schooling (Adams 2005, p. 273 and see appendix 16 there for a mapping guide to concern ISCED variables to years of schooling for different countries). The second model adds a constructed occupational status variable (HISEI) to the first model—an index of the highest household socioeconomic status constructed by the PISA team (Adams 2005, p. 273). The regression coefficients of the two models are raw inputs into the computation of two key parameters—parental education effect and total family effect—for each society. The two key parameters, estimated for each cognitive test, are the source data for the search for a simple structure of family influences across the diverse set of PISA countries.

The Latent Class Approach to Clustering

Drawing on one of the most active areas of research in statistical data analysis of survey data, this study uses latent class modeling techniques to do what statisticians have been doing with clustering techniques for much of the twentieth century. Latent class models used to be an extension of loglinear modeling of frequency tables. Over the past ten years or so, latent class models have been greatly extended to provide powerful analysis of all kinds of data and integrated into multilevel modeling frameworks. LatentGold 3.0 or above efficiently performs the estimation and classification tasks (Vermunt and Magidson. 2002).

V. Central Findings

This study is an attempt to conduct a strictly comparable cross-national analysis of the social origin of cognitive inequality, circumventing many of the problems that plagued previous research (Breen and Jonsson 2005). The empirical tests of three theoretical perspectives build on several methodological tools. The first is a simple parameterization of educational closure, simplified by the powerful identifying assumption of a relatively constant technology of within-family skill formation. The second is an application of the PSI methodology to address two fundamental problems in a comparative study of cognitive inequality. The third is the use of latent class model in the quest for any pattern of cross-national homogeneity in family effects on cognitive inequality.

In this preliminary version of the paper, I will only sketch the central findings as follows. The basic answer to the main empirical question is that there is indeed a simple structure of family effects behind seemingly large and unsystematic

cross-national variation in educational closure. This structure is effectively described by a nonparametric discrete factor model posited to account for the diverse patterns of family background effects across four cognitive outcomes and 41 countries. It is statistically adequate to describe 39 of the societies with a single nonparametric factor of just four discrete levels, two additional levels are needed to accommodate the two outlying cases—Hong Kong and Macau.[8]

The simplicity of this latent structure is reassuring. Were the methods serve nothing but introducing measurement errors or distracting from the true metric of causal relevance, the empirical analysis should not be able to capture and reveal such powerful parsimony. As in other fields of science, structural parsimony is therefore an important signal of the soundness of the methodology of this study.

More specifically, all 41 cases are remarkably consistent in their ordering along two dimensions of family effects: (1) gross parental education effect (the sum of the gross effects of paternal and maternal education), and (2) gross family effect (the sum of the effects of socioeconomic status and parental education). Figure 2 is a summary of the key findings, embedded for simplicity in a plot of average (over four cognitive outcomes) total family effect versus average gross parental education effect. The consistency of the ordering in terms of average parental education effect or average total family effect is evident. Even more striking is the consistency between the ordering along the two dimensions of effects, suggesting that parental education effects are linearly related to total family effect. The R^2 of the two dimensions for different cognitive measures are strong, ranging between .70 and .93, with an average

[8] Both societies are former colonies and now special administrative districts of China granted with a high degree of autonomy to operate their capitalist economies.

of .75.[9]

A corollary of this finding is that the two dimensions are very effective in differentiating the societies in PISA 2003. A latent class model of six clusters is able to assign membership with almost perfect accuracy. The statistical fit in terms of the BIC measure is substantially improved if the not-so-parsimonious clustering model is converted to a one-dimensional discrete factor with six ordered levels. The numerical labels of figure 2 indicate the predicted membership of a case in a four-latent-class model for four estimated coefficients of total family effect and four estimated coefficients of gross parental education effect. Table 1 lists the membership for each level. A visual inspect of the probability of membership reveals that the membership assignment is unambiguous. When a society X is assigned to level Z, the probability of X belonging to Z is almost always over 0.98 whereas the probability of X belonging to any other level is trivially small. The only two exceptions are the UK and the USA, whose probabilities of belonging to level 2 are .65 and .75, respectively, whereas their probabilities of belonging to level 3 are .35 and .25.

INSERT FIGURE 2 AND TABLE 1 ABOUT HERE.

Granted the certainty with which the model assigns membership to latent levels of the discrete factor, it is striking to find in each level a mix of societies with very different educational history and institutions. The model also groups together known cases of homogeneity, such as Czech Republic and Hungary, Norway and Sweden, Japan and Korea. Interestingly, the United Kingdom and the United States are almost quantitatively indistinguishable despite the fact that they are used to represent closed

[9] It is important to note that this strong correlation is not a mathematical necessity. Indeed, gross family effect (the algebraic sum of socioeconomic status and parental education) has R-squared of

and open societies, respectively, in stratification research. But this isomorphism is not anomalous, as it reaffirms the discovery of Kerckhoff (1974) three decades ago, when he estimates and compares the coefficients of an extended model of educational attainment for England and the United States. Even without the benefit of cross-nationally comparable estimates, Kerckhoff is so surprised by the similarity in the numerical estimates that he declares the two countries achieve the same structure and process of stratification despite the use of very different mechanisms. What he does not know, however, is the finding here that countries of very different socioeconomic development than the UK and the USA, such as Russia and Thailand, also manifest the same degree of educational closure. All these are hard to reconcile with the socioeconomic development perspective.

Ironically, post-communist states are disproportionately represented among societies with strongest family background effects. (Hong Kong and Macau, by contrast, have the lowest levels of family effects.) For instance, Hungary exhibits the strongest family effects that are much stronger than even those of the paradigmatic cases of class societies such as France. However, post-communist states are also represented in other clusters. On the other hand, Germany—a capitalist state—is represented in the cluster dominated by post-communist states. Thus the prediction of the social conflict perspective is not borne out by the evidence.

VI. Conclusion

The image of constant flux does not apply to the structure of intergenerational influence on educational achievement. As far as parental educational attainment and

---

only .06 and .61 with the status and education components, respectively. Even the stronger R-squared of .61 is substantially smaller than the R-squared with gross parental education effect.

child cognitive achievement is concerned, the image is one of highly variable flux—variable degree of intergenerational influence on educational achievement across countries. Paradoxically, against this backdrop of variable flux, there is striking isomorphism among societies of very different types of polity or levels of socioeconomic development.

Equally striking is the simple latent structure of cross-national variation in educational closure. Variation can be parsimoniously described by a single nonparametric factor with just a handful of probability masses. By and large, there are four distinct clusters of countries. Two clusters correspond to extremely high and high levels of closure in mobility. The third cluster, the modal one, is one of medium level of closure whereas a fourth cluster is distinctly open. There is some evidence for the presence of an extremely open regime of mobility, but at present there are only two data points (Hong Kong and Macau, both are tiny cities and administratively autonomous regions of China)—too few to tell whether the cases reflect a really open regime of mobility or just the ruses of exceptionally large measurement errors for the two societies.

With regard to the predictions about trends and correlations across countries, the socioeconomic development and social conflict perspectives do not get any support. The socioeconomic development perspective predicts a more or less monotonic trend toward openness or negative correlation between closure and socioeconomic development of a society. In contrast, the social conflict perspective predicts the pervasive presence of closure, especially in capitalist societies and perhaps least in Communist regimes and to some extent in post-Communist regimes. But neither perspective appears to hold. In addition, cross-national similarities in closure are not

related to the level of modernization of a society—some of the least GDP and highest GDP countries are isomorphic. Nor are the similarities related to class conflict, given of the mix of societies at both ends. In contrast, the findings are most compatible with the positional competition perspective.

This study has demonstrated the utility of a new methodological template for serious comparative analysis of cognitive inequality and, in particular, the role of parental background in the formation of cognitive differences. The methodology has wide applicability beyond the substantive questions addressed in this study. The PSI methodology is not limited to the study of family effects although it is a powerful way in this context to capture one of the sociologically meaningful aspects of family backgrounds. Nor is it limited to the study of cognitive inequality on the dependent variable side.

References

Adams, Ray. 2005. *PISA 2003 Technical Report: Programme for International Student Assessment*. Paris: Organisation for Economic Co-operation and Development.

Blau, Peter M. and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Boudon, Raymond. 1974. *Education, Opportunity and Social Inequality*. NY: Wiley.

Bourdieu, Pierre. 1973. "Cultural Reproduction and Social Reproduction." Pp. 71-112 in *Knowledge, Education, and Cultural Change*, edited by R. Brown. London: Tavistock.

Breen, Richard, and Jan O. Jonsson. 2005. "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility." *Annual Review of Sociology* 31:223–243.

Cameron, Stephen V., and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy* 106:262–333.

Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109:455–99.

Carneiro, Pedro, and James J. Heckman. 2003. "Human Capital Policy." Pp. 77-239 in *Inequality in America: What Role for Human Capital Policies?*, edited by J. J. Heckman, A. B. Krueger, and B. M. Friedman. Cambridge, MA: MIT.

Coleman, James S., and others. 1966. *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.

Collins, Randall. 1971. "Functional and Conflict Theories of Educational Stratification." *American Sociological Review* 36(6): 1002-1019.

Collins, Randall. 1979. *The Credential Society*. New York: Academic Press.

Erikson, Robert, and John H. Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.

Featherman, David L. and Robert M. Hauser. 1978. *Opportunity and change*. New York: Academic Press.

Frank, Robert H. 1985. *Choosing the Right Pond: Human Behavior and the Quest for Status*. New York: Oxford University Press.

Hanushek, Eric A. and John E. Jackson. 1977. *Statistical Methods for Social Scientists*. San Diego: Academic Press.

Hauser, Robert M. and David L. Featherman. 1977. *The Process of Stratification: Trends and Analyses*. New York: Academic Press.

Heckman, James J., Alan B. Krueger, and Benjamin M. Friedman. 2003. *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT.

Herrnstein, Richard J., and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

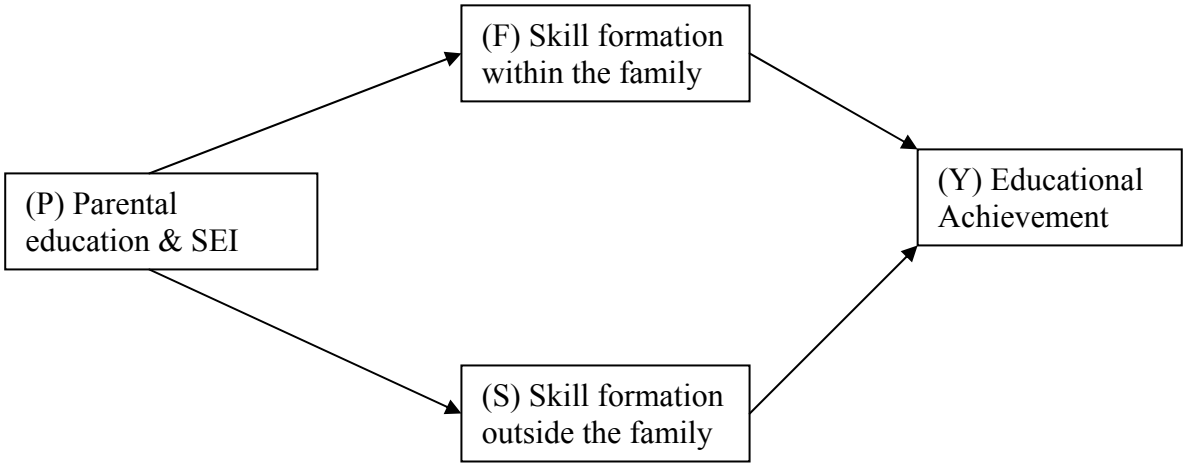
Hirsch, Fred. 1976. *Social Limits to Growth*. Cambridge, Mass.: Harvard University Press.

Hout, Michael, and Tom DiPrete. 2006. "What We Have Learned: RC28's Contributions to Knowledge about Social Stratification." *Research in Social Stratification and Mobility* 24:1-20.

Jencks, Christopher. 1972. Inequality: A Reassessment of the Effect of Family and

Schooling in America. New York: Basic Books.
Jencks, Christopher, et al. 1979. *Who Gets Ahead*. New York: Basic Books.
Kerckhoff, Alan C. 1974. "Stratification Processes and Outcomes in England and the U.S." *American Sociological Review* 39:789-801.
Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89:841-864.
Lieberson, Stanley. 1985. *Making It Count*. Berkeley and Los Angeles, CA: University of California Press.
Lord, F. M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Erlbaum.
Parsons, Talcott. 1970. "Equality and Inequality in Modern Society, or Social Stratification Revisited." *Social Stratification*, edited by Edward O. Laumann. Indianapolis: Bobbs-Merrill.
Rasch, G. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: University of Chicago Press.
Sewell, William H., and Robert M. Hauser. 1975. *Education, Occupation and Earnings*. NY: Academic Press.
Sorensen, Aage B. 1977. "The Structure of Inequality and the Process of Attainment." *American Sociological Review* 42:965-978.
Tam, Tony. 2005. "Comparing the Seemingly Incomparable: A Theoretical Method for the Analysis of Cognitive Inequality." Paper presented at the conference on Welfare States and Social Inequality, RC28 of the International Sociological Association, Oslo, Norway, May 5-8, 2005.
Thurow, Lester C. 1975. *Generating Inequality*. London: MacMillan.
Treiman, Donald J. 1970. "Industrialization and Social Stratification". *Social Stratification*, edited by Edward O. Laumann. Indianapolis: Bobbs-Merrill.
Vermunt, Jeroen, and Jay Magidson. 2002. "Latent Class Cluster Analysis." Pp. 89-106 in *Applied Latent Class Analysis*, edited by J. A. Hagenaars and A. L. McCutcheon. Cambridge: Cambridge University Press.
White, Harrison C. 1970. *Chains of Opportunity*. Cambridge: Harvard University Press.

A. The Basic Framework



B. A Path-analytic Representation

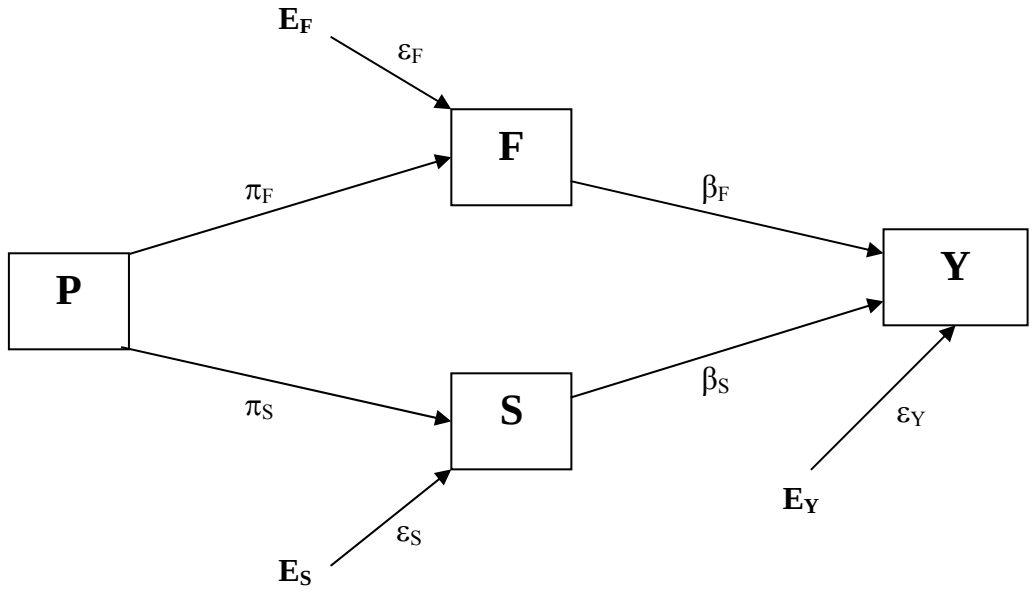


Figure 1. A conceptual framework for defining and comparing intergenerational educational mobility across societies and over time. Heuristically, $\pi_F\beta_F$ reflects what researchers would call "direct" effects of parental background because the upper path **P-F-Y** is based on infra-family processes. Similarly, $\pi_S\beta_S$ reflects what may be called "indirect" effects of parental background, indirect because the lower path **P-S-Y** is mediated by formal and informal schooling institutions outside the family and over the entire schooling life cycle.

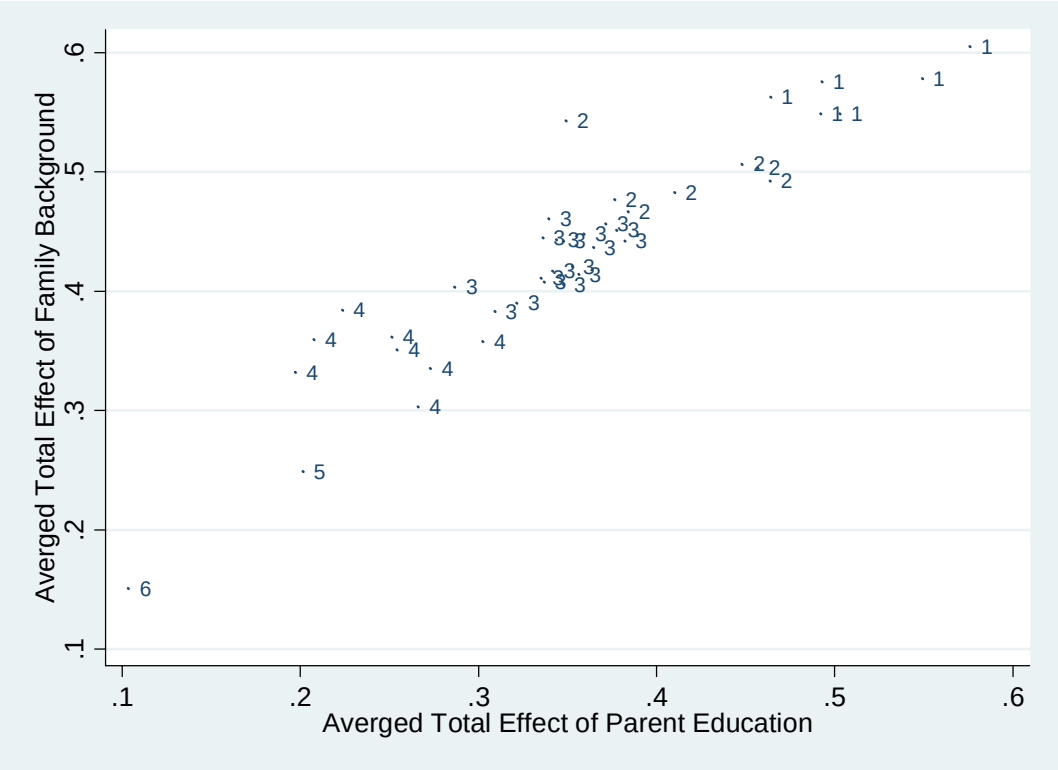


Figure 2. The distribution of total family effect and total parental education effect on four dimensions of cognitive outcomes across 41 countries in PISA 2003.

Table 1. Forty-one PISA 2003 Countries Listed by the Levels of Family Background Effects on Cognitive Achievement Status

Highest Level 1 N=6	**Level 2 N=7**	**Level 3 N=18**		**Level 4 N=8**	**Level 5 N=1**	**Lowest Level 6 N=1**
Czech Republic	France	Australia	Luxembourg	Brazil	Hong Kong	Macao
Germany	Liechtenstein	Austria	Mexico	Finland		
Hungary	Poland	Belgium	Netherlands	Iceland		
Slovak Republic	Russia	Canada	New Zealand	Indonesia		
Turkey	Thailand	Denmark	Portugal	Italy		
Uruguay	UK	Greece	Spain	Latvia		
	USA	Ireland	Switzerland	Norway		
		Japan	Tunisia	Sweden		
		Korea, South	Yugoslavia, Former			