University of Toronto
Department of Economics

Working Paper 552

Evaluating Progress Toward an Equal Opportunity Goal: Assessing the German Educational Reforms of the First Decade of the 21st Century.

By Gordon Anderson, Thomas Fruehauf, Maria Grazia Pittau and Roberto Zelli

November 20, 2015

Evaluating Progress Toward an Equal Opportunity Goal: Assessing the German Educational Reforms of the First Decade of the 21^{st} Century.

Gordon Anderson, Thomas Fruehauf*
Maria Grazia Pittau and Roberto Zelli†

November 20, 2015

Abstract

The first Program for International Student Assessment (PISA) results on language literacy administered in Germany in 2000 shocked the nation bringing about some fundamental reforms in the education system. A €4 billion plan to reform the schooling system involved intensified parent and teacher training, increases in the number of schooling hours and changes in the way student performance was evaluated. By way of measuring the extent of the improvements, this paper proposes and implements new techniques for evaluating the effectiveness of the reforms in the context of a social justice imperative when outcomes before and after their introduction are not cardinally comparable. Fundamental changes in the structure of the dependency of child outcomes on circumstances were detected with some qualified improvements in equality of opportunity over the period.

Keywords: Educational attainment, Equality of opportunity, PISA program.
JEL classification: I21, C14.

*Department of Economics, University of Toronto,
email: `anderson@chass.utoronto.ca`, `thomasfruehauf@yahoo.ca`
†Department of Statistical Sciences, Sapienza University of Rome,
email: `grazia.pittau@uniroma1.it`, `roberto.zelli@uniroma1.it`.

1 Introduction.

The first Program for International Student Assessment (PISA) results on language literacy administered in Germany in 2000 shocked the nation, the country came well below the average overall for all the countries tested and it did no better in mathematics and science than it did in language. The correlation between its family socio-economic status and student achievement was higher than any other OECD country undermining a long held view in Germany that the choice of secondary school is based solely on achievement in elementary school. It has been observed that even when students were matched on actual achievement, elementary school children whose parents had attended the highest school level were three times as likely to be sent to that same highest level of school as children of parents who had graduated from the lowest school level (OECD, 2011). In short there was overwhelming evidence of a dependence of a child's educational outcomes on the circumstances they confronted, an evident lack of equality of opportunity.

These results prompted a €4 billion plan to reform the schooling system, improve student outcomes and equalize opportunities (OECD, 2011). The reforms included more involvement in Kindergarten, in teaching German to immigrant children, more funds for special language training to help non-German speaking families, more funds for all-day school systems, and more funds for teacher training programs. In 2003–2004, national education standards were put in place (and subsequently raised in 2007) and over the following years, the German States adjusted their testing procedures to more closely correspond to PISA test standards and came together to harmonize their curricula, create improved tests and raise the bar yet again. Such policies could not only elevate the outcomes of students in particular groups but they could also change the nature and number of the particular groupings. The question is how to measure the success of the reforms in this particular context when test outcomes before and after the reforms cannot be cardinally compared because of changes in curriculum, standards and testing procedures. In essence the difficulty is that average or median test scores before and after the reforms are not comparable because of changes in the testing methodologies so that increases (or decreases) in such statistics are not indications of improvements (or otherwise) resulting from the policy. However it may be possible to assess whether or not the policy has brought about child educational attainments in a more "socially just" fashion.

Assessment of the degree of social justice is difficult, in the context of education it usually relates to the notion of equality of opportunity (Roemer 1998, 2006, 2010). Atkinson (2012) and Sen (2009) argue that the aim is to seek progressive reform rather than transcendental optimality. Accordingly, techniques for evaluating such progress should be capable of measuring the degree and significance of such advances or retreats in economic and social outcomes. The social desire for equality of opportunity is the requirement that an individuals' achievement should depend upon their effort and choice and not be predicated upon the circumstances they face for which they cannot be held responsible. Policy can thus be directed toward equalizing outcomes of individuals with the same effort and choices, or equalizing unambiguous inequalities in circumstances. Policy in Germany clearly took both directions in investing in language education for immigrant parents (essentially elevating the circumstances of their children), in supplemental language education for immigrant children as well as substantive reforms elsewhere in the educational system. Such policies may have not only elevated the outcomes of particular classes or groups of people, but they could also have changed the numbers of groups by eliminating or creating classes. In order to evaluate the impact of such policies, measures of the extent to which the structure of the relationship between effort, outcomes and circumstances have changed in terms of both types and numbers of classes are required.

One major difficulty with such measurement is that much in the nature of individual choice, effort and circumstance is fundamentally multidimensional and unobservable. The vast literature on functionings and capabilities suggests that it is quite possible for people with the same choices and circumstances to have very different achievements (because of variation in effort) that are only partially related to their circumstance differences. Similarly observable circumstances are not deterministically related to their fundamental circumstances but more generally they are drawn from a circumstance class (for example genetic endowments are not an inconsequential component of a child's circumstances as is the level of nurturing of the child which is really a matter of parental effort). This inevitably involves the classification of individual achievements and circumstances into classes so that in an analytic context, it is the extent to which membership of an achievement class is predicated upon membership of a given circumstance class that is of interest.

The long tradition in the empirical economics literature of classifying agents into groups has invariably involved specifying “arbitrary” boundaries (for example income quantiles in the income mobility literature, high school grade levels in the educational literature and the poverty frontier in the poorness literature) for set inclusion and exclusion purposes which have been a matter of much debate[1]. Recently Fituossi, Sen and Stiglitz (2011) summarize a literature which argued that poorness and wellness are many dimensioned concepts - partitioning an income distribution is but a vague reflection of the true boundary. In addition Sen and others (Grusky and Kanbur, 2006; Kakwani and Silber, 2008; Nussbaum, 2011; Alkire and Foster, 2011) have forcibly argued that the fundamentally unobservable or hard to quantify limitations to people’s functionings and capabilities are the determining or bounding factors in individual achievements. Problem: when the determinants of class become multidimensional and in part unobservable, boundaries become much more difficult to define so that agents with the same achievements may come from ostensibly different classes (Anderson *et* al., 2015), essentially, in the context of an achievement distribution, the number of classes and their boundaries become fuzzy at best.

In the development here achievement is associated with latent effort/choice variables which may or may not be influenced by partially observable circumstances. The transition from circumstance classes to achievement classes over which the effort of agents is assumed to be normally distributed is studied. A technique for “partially” determining class membership in the presence of a latent effort/choice environment is proposed which avoids the arbitrary specification of the number of classes or their boundaries. If these many unobservable, hard to measure factors determine peoples observable behavior and people within a particular class face similar limits (which differ across classes), it may be possible to partially discern distinct individual classes and representations of behaviors common to a particular classification. Class boundaries are “partially determined” in the sense that all that is established is the probability that a child with a given set of grades is in a particular class, but this is shown not

[1]See for example Atkinson and Brandolini, 2013; Banerjee and Duflo, 2008; Beach, Chaykovski and Slotsve, 1997; Citro and Michael, 1995; Easterly, 2001; Foster, 1998; Milanovic and Yitzhaki, 2002; Saez and Veall, 2005; Sen, 1983; Townsend, 1985.

to hinder the study of the number of classes and individual class behavior nor the statistical relationship between outcome classes and circumstance classes. Furthermore, unlike standard transitional approaches, the technique admits the possibility that the achievement classes and the circumstance classes may differ in number. Possibilities for evaluating the progress of equality of opportunity using these techniques are explored using the PISA data set for Germany for the years 2003 and 2009 i.e. the year before the reforms were introduced and the year sometime after the reforms were introduced.

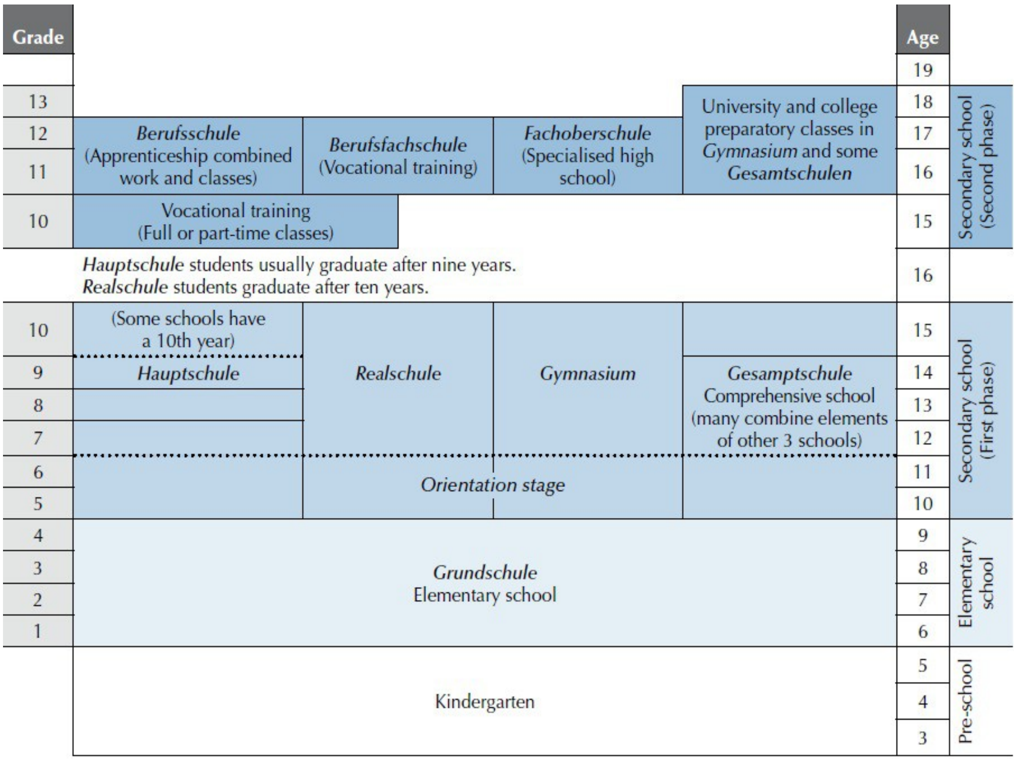
In the following the background context to the educational reforms in Germany is explored in Section 2. Section 3 introduces the basic model and the approach to determining outcome and circumstance classes and develops the tools for analysis of the extent to which Equality of Opportunity has progressed over the period. Section 4 reports the results and Section 5 draws some conclusions.

2 Background

Throughout the latter half of the 20th century Germany had uniform elementary schooling with compulsory elementary education for all children aged 6 through 10. On completion students were streamed into one of three types of school, Hauptschule, Realschule or Gymnasium. The vast majority of students of lowest abilities were streamed into Hauptschule where, after a few more years of education, they would receive a qualification entitling them to apply for training generally leading to low skilled blue collar jobs. This training was often obtained after grade 9 between the ages 15 and 17 through vocational training at a Berufsschule, a combination of apprenticeship and classes. More qualified students enrolled into the Realschule. These students would generally attend school for at least one extra year after which they got vocational training at a Berufsfachschule, tailored to a more specific training of clerks, technicians and lower-level civil servants.[2] The highest secondary school was the Gymnasium. It focused on broad preparation in the humanities and the Abitur, which was the sole gateway to the professions, teaching and the upper levels of the

[2]This secondary schooling system was augmented by what is referred to as the "Dual System" managed at the federal level whereby students who complete a secondary education can be apprenticed to a firm whilst simultaneously attending a vocational "continuation school". Until recently, the Hauptschule was the primary source of employee's apprentices with some white collar jobs going to students from the Realschule.

Figure 1: *Breakdown of Germany's educational system*



civil service. These students completed preparatory classes for university and college and attended Fachoberschule after grade 10, a specialized high school. Figure 1 shows a breakdown of Germany's educational system.

These school divisions seem to reflect the social divisions in Germany. The Gymnasium was for the offspring of nobility and the upper middle classes, Realschule for the offspring of the lower middle class citizens and Hauptschule for the working class, which is where most employees got their apprentices. The Gast Arbeiter program solution to the labour shortages of the latter quarter of the 20th century saw a considerable influx of workers from countries with comparatively low levels of education. Many settled and raised families and those successive generations typically had poor German language skills. A by-product of the Gast Arbeiter program was that at a certain point the immigrant population made up more than half of the students in elementary schools in some Northern German cities. Increasing demands for high skilled workers led increasing demand for entry into the Gymnasium and Realschule and consequently the Hauptschule was confined to students with limited perspectives, that include immigrants and native German children from lower class families.

In response to the concerns raised by the 2000 PISA results several programs were initiated upon to elevate standards and reduce the influence of socio economic background and immigrant status, often associated with lower German language skills, on student outcomes. An all-day school program was embarked upon in 2003 with substantial funding to extend the school day until 4:00 pm or later. Prior to this increase, German students aged 7 and 8 spent an average of 626 hours a year in school, compared to 788 in the OECD countries. There was investment in more teacher training to help diagnose and address specific problems faced by students with needs for learning aids and supplementary language tuition. The reform brought fundamental change to the old feudal structure of the schools, which had focused on having few highly educated people, several with medium education and the majority with little education. Kindergarten quality was enhanced with special language training for families with little knowledge of German language. The German educational system reserved a place in Kindergarten for every child from the age of 3 until they begin elementary school to help improve their German language skills. Educators became more accountable for their performance and in exchange they were given more autonomy to achieve higher testing standards demanded by PISA tests.

In 2003 and 2004 national educational standards were introduced for children in primary and secondary school in German language, mathematics, a first foreign language (English or French), and science (biology, chemistry and physics). Additional standards were put in place for students at the end of grade 10 in 2007. These performance standards covered subject-specific competencies at a similar level as the PISA tests that students were expected to meet throughout Germany. Prior to this now mandatory agreement, there had never been national standards in Germany.

In 2006, the Council of ministers agreed to develop common assessments based on a national scale for 3rd graders in elementary school, and 8th and 9th graders in secondary schools in all 16 German states. While these assessments did not carry high stakes, the sample was still representative in each state. To make the yearly spring tests more comprehensive some states joined forces to develop testing systems that obey the new curriculum standards.

Implementation of these new policies was helped by the Institute for Educational Progress (Institut zur Qualittäsentwicklung im Bildungswesen (IQB)) based at the Humboldt University in Berlin. It monitored the new education system by collecting and analyzing information in order to check performance of students and teachers. Germany also took part in the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS) at the elementary school level to continuously assess the competency level of its students.

To increase the chance of a student to attend a higher secondary school some states delayed the separation into the tripartite system, Hauptschule, Realschule or Gymnasium, until the student was 12 rather than 10 years old. Other states combined the Realschule and Haptschule into one school, while some allowed students in lower schools to move up the ladder and complete their education with a more prestigious background, allowing for better job-opportunities. This led to speculations in 2008/2009 that the 2,625 Realschulen and 4,283 Hauptschulen will no longer co-exist within 10 years and will merge into one type of school. There were 3,070 Gymnasien during that time, less than half of the other two school types combined.

3 The Model and Tools for Evaluating Transitional States

PISA data for German students in 2003 (before the reforms had taken effect) and 2009 (after the reforms had been implemented) is used to construct an achievement index for students who have completed exams in Math ($X1$), Language ($X2$), and Science ($X3$). Because the evaluation methodology in the different disciplines differed slightly over the two observation periods the overall achievement index was based upon the average of their maximum mark standardized subject scores i.e. for person i:

$$\text{Achievement Score}(i) = \frac{1}{3} \cdot \left\{ \frac{X1_i}{\text{Max}(X1)} + \frac{X2_i}{\text{Max}(X2)} + \frac{X3_i}{\text{Max}(X3)} \right\}$$

To develop circumstance classes that reflect a student's parental environment (circumstances) we construct an index by adding the educational status (a six point scale) of each parent present in the household and divide by the square root of the number of parents present. This is akin to using the square root rule for parental circumstance support common in consumer equivalence scaling (Brady and Barber, 1948) wherein there is an advantage to the presence of more than one parent but it is an advantage with diminishing returns to scale (0.5 elasticity). This index is then used to define three circumstance categories: Lower, Middle and Upper of roughly equal sizes in the initial year by exploiting gaps in the index scale so that, unlike the achievement variable circumstance class membership is definitively discrete.

Given that observable student outcomes (aggregate achievement scores) are governed by three unobservable factors, their innate abilities, choices and effort, it is assumed that there are a finite number K of student achievement types or classes labeled $k = 1, \cdots, K$ where "k type" students, have similar abilities and choices, effort is however randomly distributed across the same type. Following Gibrat's law (Sutton, 1997) we suppose that x_{tk}, the achievement score for a student in achievement class k in period t, follows the law of proportionate effects with v_{tk} its outcome improvement rate in period t, and let T be the elapsed time period over which the student has progressed with x_{0k} the initial achievement level. Assuming the $v's$ to be independent and identically distributed random variables with a small (relative to one) mean δ_k and finite variance σ_k^2 which vary with type, it may be shown that for an elapsed schooling period of T, the log achievement size distribution of such students would be linked systematically from period to period in terms of means and variances in the form:

$$\ln(x_{Tk}) \sim \text{N}(\ln(x_{0k}) + T(\delta_k + 0.5\sigma_k^2), T\sigma_k^2).$$

Note that the distribution is governed by the initial condition $\ln(x_{0k})$ and the growth rate δ_k which in turn are dependent on the circumstances, innate abilities as well as efforts of the student. This does imply that the achievement distributions of different classes overlap so that the achievement of a high effort student from a low achievement class could exceed that of a low effort student from a higher achievement class. The size distribution of achievements of a collection of students will be a weighted sum of these achievement class distributions where the weights equal the proportions of the student population in the corresponding classes. The unobservable factors are, to some degree, influenced by a student's partially observed circumstances, the nature and nurture effects of social and parental background, thus it is assumed that the chance that a student is in a particular achievement class is partially determined by her circumstances. Given the probability that a student with achievement x is in class k and the knowledge that she is from a particular circumstance class facilitates study of the relationship between achievements and circumstances.

3.1 Mixture Models and Class Membership Probabilities

For generality purposes, suppose K achievement classes emanating from J circumstance classes are contemplated. Achievements are continuously measured on the unit interval, while circumstances are discrete ordered categories. For a given achievement class k the achievement of student i, x_i, may be approximately written as $x_i = \mu_k + \sigma_k \cdot e_i$ where $e_i \sim \mathrm{N}(0,1)$, so that $\sigma_k \cdot e_i$ is a latent measure of student i's effort. Thus, the distribution of x is given by:

$$f(x, \mathbf{\Psi}) = \sum_{k=1}^{K} w_k f_k(x, \theta_k), \tag{1}$$

where $f_k(x, \theta_k)$ are $\mathrm{N}(\mu_k, \sigma_k)$.[3] The vector $\mathbf{\Psi} = (w_1, \cdots, w_{K-1}, \xi')'$ contains all the unknown parameters of the mixture model: w_k, $k = 1, \cdots, K$ are the mixing proportions summing to 1 ($\sum_{k=1}^{K} w_k = 1$); the vector ξ contains all the parameters $(\theta_1, \cdots, \theta_K)$ known *a priori* to be distinct. The w_k represent the *a priori* probabilities of a randomly selected agent in the population to belong to achievement class k. They are endogenous parameters which determine the relative importance of each component in the mixture and can be interpreted as unconditional probabilities. By a simple limiting argument (Anderson *et* al., 2015), the conditional probability of an agent i

[3]The choice of normal densities depend on the assumption of normality in effort. However this is not an overly strong assumption since, any continuous distribution can be approximated to some desired degree of accuracy by an appropriate finite Gaussian mixture (Marron and Wand, 1992; Rossi, 2014)

$(i = 1, \cdots, n)$ with achievement x_i being in achievement class k $(k = 1, \cdots, K)$ is given by:

$$\pi_{ik}^A = \texttt{Prob}\{A(i) = k \mid (x_i; \mathbf{\Psi})\} = \frac{w_k f_k(x_i)}{\sum_{k=1}^K w_k f_k(x_i)} \tag{2}$$

where $A(i)$ indicates the achievement class component to which agent i belongs, yielding the probability of achievement for each agent i to belong to the mixture component k.

In estimating the parameters, the class weights (the unconditional probabilities), are estimated by using the individual class weights π_{ik}^A as:

$$\widehat{\pi}_k^A = \widehat{w}_k = \frac{1}{n} \sum_{i=1}^n \pi_{ik}^A, \;\; k = 1, \cdots, K. \tag{3}$$

Given the number of classes K, the unknown parameters of the mixture (means, variances and proportions of each component) along with the conditional probabilities (π_{ik}^A) are estimated by maximum likelihood (ML) via the expectation-maximization (EM) algorithm (Dempster *et al.*, 1977). Starting from a given number of components and an initial parameter $\mathbf{\Psi}^{(0)}$, the first stage of the algorithm (E-step) is to assign to each data point its current conditional probabilities. In the second stage (M-step), the maximum likelihood estimates are computed using the conditional probabilities as conditional mixing weights. The estimates of the parameters are used to re-attribute a set of improved probabilities of group membership and the sequence of alternate E and M steps continues until a satisfactory degree of convergence occurs to the ML estimates.[4]

The probability of a randomly selected agent to belong to a given circumstance class j $(j = 1, \cdots, J)$ is π_j^c. Given a sample $i = 1, \cdots, n$, π_j^c is estimated as:

$$\widehat{\pi}_j^c = \frac{1}{n} \sum_{i=1}^n D_{ij} \tag{4}$$

[4]It is well known that the likelihood function of normal mixtures is unbounded and the global maximizer does not exist (McLachlan and Peel, 2000). Therefore, the maximum likelihood estimator of $\mathbf{\Psi}$ should be the root of the likelihood equation corresponding to the largest of the local maxima located. The solution usually adopted is to apply a range of starting solutions for the iterations. In this paper, randomly selected starts, large sample non-hierarchical (Kaufman and Rousseeuw, 1990) clustering-based starts have been selected for initialization.

where:

$$D_{ij} = \begin{cases} 1 & \text{if agent } i \text{ has circumstance } j \\ 0 & \text{otherwise} \end{cases}$$

Let the $K \times J$ matrix T whose typical element is $t_{kj}(k = 1, \cdots, K; j = 1, \cdots, J)$ be the transition matrix yielding the conditional probability of being in achievement class k given circumstance class j. The $K \times 1$ vector of achievement class probabilities π^A whose typical element is π_k^A is related to the circumstance class probability vector π^C (whose typical element is π_j^C) by the formula:

$$\pi^A = T \cdot \pi^C \tag{5}$$

Given a sample of agents, matrix T may be estimated by a simple regression system of the form:

$$\pi_i^A = T \cdot D_i + \nu_i, \ \ i = 1, \cdots, n \tag{6}$$

where π_i^A is the $K \times 1$ vector of the conditional probabilities of agent i, whose typical element is π_{ik}^A; D_i is the $J \times 1$ vector whose typical element is D_{ij} and ν_i is a $K \times 1$ random vector with zero mean and singular covariance matrix. Thus $\pi^A = \mathrm{E}\left(\pi_i^A\right) = \mathrm{E}\left(TD_i + \nu_i\right) = T \cdot \mathrm{E}\left(D_i\right) = T \cdot \pi^C$, where the columns of T sum to 1. The resulting estimated probabilities are:

$$\widehat{T}_{kj} = \frac{1}{|C(i) = j|} \sum_{i \in C(i) = j} \pi_{ik}^A.$$

Estimating T can be seen to be a matter of estimating $M\mathrm{diag}(\pi^C)^{-1}$, where M is the joint probability distribution of achievements and circumstances and $\mathrm{diag}(\pi^C)$ is a matrix with circumstance class probabilities π^C on the diagonal, so that M can be retrieved by post multiplying T by $\mathrm{diag}(\pi^C)$ or estimated directly by multiplying each of the circumstance dummies by the inverse of the probability of being in that circumstance class.

3.2 Choosing the number of achievement classes.

Selection of K for the achievement distribution is performed by measuring the proximity of the mixture distribution, $f(x, \mathbf{\Psi})$, to a kernel estimate of the distribution,

$f_{krn}(x)$, of achievements by using two versions of Gini's Transvariation Coefficient as proposed in Gini (1916), which measures the dissimilarity of two distributions, modified by a penalty factor. Following arguments in Akaike (1972), the penalty is the number of coefficients in the mixture times $2/n$ where n is the sample size.

The two versions of Gini's Transvariation Coefficient, GTR and GTRIM, a weighted version of GTR, relate to the integral of absolute differences between two probability distribution functions.[5] Particularly, GTR refers to the following overlap measure: $\theta = \int_{-\infty}^{\infty} \min \{f(x, \mathbf{\Psi}), f_{krn}(x)\} \, dx$:

$$\text{GTR} = \int_{-\infty}^{\infty} |f(x, \mathbf{\Psi}) - f_{krn}(x)| dx = 2 - 2\theta \tag{7}$$

Anderson, Linton and Whang (2013) showed the overlap estimator $\widehat{\theta}$ to be asymptotically normally distributed with mean equal to θ and a certain variance V, and therefore GTR $\sim \text{N}(2 - 2\theta, 4V)$, thus facilitating inference for GTR.

GTRIM is an importance weighted version of GTR:

$$\text{GTRIM} = \int_{-\infty}^{\infty} |f(x, \mathbf{\Psi}) - f_{krn}(x)| f_{krn}^{-0.5}(x) dx \tag{8}$$

Gini's transvariation coefficient can be seen as cumulating the absolute difference between the functions over the whole real line, whereas the GTRIM version can be seen as cumulating the "importance" weighted absolute difference. It can be seen as weighting the difference of $(x, \mathbf{\Psi})$ from the "target" $f_{krn}(x)$ by some monotonic function of the "target" function, so that a given difference from a small target plays a bigger role in the calculation than the same order of difference in a correspondingly larger target. In essence, the differences are weighted with respect to some reference distribution. Choice of the square root function is very much inspired by, and in the spirit of, entropic measures of variation such as Theil, Pearson and Shannon. For example the continuous version of Theil's entropic measure (TE) is related to the continuous version of Pearsons Chi squared (PCHI) dissimilarity measure as follows:

$$\text{TE} = \int f(x, \mathbf{\Psi}) \ln \left(\frac{f(x, \mathbf{\Psi})}{f_{krn}(x)} \right) dx = \int f(x, \mathbf{\Psi}) \ln \left(1 + \frac{f(x, \mathbf{\Psi}) - f_{krn}(x)}{f_{krn}(x)} \right) dx$$

$$\approx \int f(x, \mathbf{\Psi}) \left(\frac{f(x, \mathbf{\Psi}) - f_{krn}(x)}{f_{krn}(x)} \right) dx = \int \left(\frac{[f(x, \mathbf{\Psi}) - f_{krn}(x)]^2}{f_{krn}(x)} \right) dx = \text{PCHI}$$

[5] Assume for convenience that $f_{krn}(x)$ has positive support over the whole real line.

The argument under the integral sign in GTRIM can be seen to be the square root of the argument under the integral sign of PCHI.

The optimal number of components in the mixture is consequently assessed comparing the mixture distribution with the true unknown density, consistently estimated by a kernel estimator. Namely, this is picking the value K that minimizes the penalized GTR or GTRIM.

3.3 Measuring the Extent of and Changes in Equality of Opportunity: New Mobility Indices and Methods for Evaluating Transitions

Since this framework allows for non-square transition matrices (the numbers of circumstance J_C and achievement classes K_A are not necessarily the same), the Shorrocks (1978) suggestion of a simple index of mobility (equality of opportunity) as $(K - \text{trace}(T))/(K-1)$, where K is the number of categories or dimension of T is not viable. When achievements are independent of circumstances, the joint probability matrix $M^* = \pi^A \cdot (\pi^C)'$ and $T^* = M^* \text{diag}(\pi^C)^{-1}$. An index of general equality of opportunity (Anderson and Leo, 2015) is afforded by the degree of overlap between T and T^*:

$$\text{OVLP} = \sum_{k=1}^{K_A} \sum_{j=1}^{J_c} \min\left(T_{kj}, T^*_{kj}\right)$$

However, instead of estimating M, we can proceed as follows. The jth column of T corresponds to the probability distribution over the final state outcome space for agents emerging from initial state j. As such, perfect mobility (where the final state is uninfluenced by or independent of the initial state) is characterized by T having common columns which all sum to 1. Writing the kth row of T as t_k, let `maxr()` and `minr()` be operators which return the maximum and minimum value in a row vector respectively, an index of mobility, (TM), which immediately suggests itself is:

$$\text{TM}(T) = 1 - \frac{\sum_{k=1}^{K_A} \left(\texttt{maxr}(t_k) - \texttt{minr}(t_k)\right)}{\min\left(J_C, K_A\right)} \tag{9}$$

The index TM can be viewed as a multivariate scaled version of Gini's two distribution dissimilarity "transvariation" index (Gini, 1916). The index $TM(T)$ ranges from 0 to 1. In case of perfect mobility each column of T will be identical, the final state outcome distributions emerging from the J_C initial states will overlap perfectly and therefore the sum of maximums will equal the sum of minimums, yielding $TM = 1$.

In case of perfect immobility, we have two different situations.

1. The number of achievement classes is less or equal to the number of circumstances: $K_A \leq J_C$. The columns of T are orthogonal, the final state outcome distributions do not overlap, the sum of minimums will be 0 and the sum of maximums will be equal to J_C, the the number of conditional distributions, yielding $TM(T) = 0$. In the context of square transition matrices complete immobility is characterized by $T = I$, the identity matrix.

2. The number of achievement classes is greater than the number of circumstances: $K_A > J_C$. In this case only K_A columns of T are different and orthogonal, while the remaining $J_C - K_A$ columns are identical to the previous ones. The sum of minimums will be 0 and the sum of maximums will be equal to K_A, the number of independent conditional distributions, yielding $TM(T) = 0$.

When T is a square matrix, $TM(T)$ satisfies the normalization, immobility and perfect mobility axioms of Shorrocks (1978). It also satisfies the strong perfect mobility axiom since $TM(T) = 1$ if and only if T has common columns. However, it does not satisfy strong perfect immobility axiom (that is $TM(T) = 0$ if and only if $T = I$) since $TM(T) = 0$ for any column rearrangement of the identity matrix. The monotonicity axiom that requires $TM(T) > TM(\breve{T})$ when $T_{kj} \geq \breve{T}_{kj}$ for all $k \neq j$ with strict inequality holding somewhere, is satisfied[6]. Period consistency requires $TM(T) \geq TM(\breve{T})$ implies $TM(T^s) \geq TM(\breve{T}^s)$ for positive integer $s > 0$. Finally, when the outcome and circumstance variables only have an ordinal ranking (i.e. they cannot be cardinally compared) the index can be shown to have the property of scale invariance and scale independence (see for example Kobus and Milos, 2012).

It has been suggested that the achievement of equality of opportunity can be assessed by establishing the absence of dominance relationships between the outcome distributions of circumstances (Lefranc *et* al., 2009; Dardanoni, 1993) but, in the absence of the achievement of the equal opportunity goal, little can be gleaned from this

[6]The incremental increase in any off diagonal element requires a concomitant decrease in its corresponding column on diagonal element (to preserve adding up). Such a change can only decrease $sum(maxr(t) - minr(t))$ and hence increase $TM(T)$

approach. Anderson and Leo (2015) propose an index, which is asymptotically normally distributed with an estimable standard error, based upon the area between the dominance curves at a given order as a measure of distance from an equal opportunity goal with reductions in the index, indicating progress toward equality of opportunity. Thus, for example at the Mth order dominance level, we contemplate:

$$\int_0^1 \int_0^{x_M} \cdots \int_0^{x_1} \sum_{k=1}^{K} w_k f_k \left(x|\text{Low parental class}\right) - \\ \sum_{k=1}^{K} w_k f_k \left(x|\text{High parental class}\right) dx dx_1 \cdots dx_M$$

In essence this is, at the i'th order of dominance, a measure of how far apart are the outcomes of the poorest parentally endowed children and the richest parentally endowed children.[7]

3.4 Viewing the circumstance–achievement transition as a process

In the context of a generational model, where this generation's achievement classes become the circumstance classes of a subsequent generation, the transition matrix can be viewed as characterizing a process. Anderson (2015) demonstrates how the transition matrix can be used to evaluate whether the transitions are converging or polarizing the outcome or achievement distribution via a balance of probabilities measure. In the present context a convergent transition matrix implies equalizing the circumstances for subsequent generations.

For expositional simplicity suppose the transition matrix is aggregated into a 3×3 matrix with typical element T_{kj} $k, j = 1, 2, 3$ of transitions to low, middle and high achievements from low, middle, high circumstances, with circumstance probabilities π_j^c $j = 1, 2, 3$.[8] Index PT is defined as the probability of an agent with non-middle class circumstances achieving middle class outcomes less the probability of an agent with middle class circumstances achieving non middle class outcomes. when PT is positive, a convergent process is indicated (polarizing when it is negative). A similar balance

[7] Unfortunately since, because of different testing methodologies, outcome distributions in the two observation years are not strictly commensurable only within year comparisons can be made.

[8] Given more than 3 outcome and or circumstance groups it is a simple task to aggregate the transition matrix into a 3×3 grouping.

of probability measure, PUT, can assess mobility as upward or downward transiting. From an Equality of Opportunity perspective convergent processes are to be preferred to polarizing processes since they may be seen as equalizing the circumstances of subsequent generations and upward transiting rather than downward transiting processes are to be preferred since they are improving the circumstances of future generations. These statistics may be written as:

$$PT = wT_{21} + (1 - w)T_{23} - (T_{12} + T_{32}), \text{ where } w = \frac{\pi_1^c}{\pi_1^c + \pi_3^c};$$

$$PUT = (1 - T_{11})\pi_1^c - (1 - T_{33})\pi_3^c - (T_{12} - T_{32})\pi_2^c$$

.

4 Empirical evidence

The Program for International Student Assessment (PISA) results for Germany in the years 2003 and 2009 were employed in this study. Table 1 reports summary statistics of the raw data and the constructed achievement and circumstance variables to be used in this study.

Table 1: *Summary Statistics*

	2003 (n=832)					2009 (n=1627)				
	max	min	mean	median	std.dev.	max	min	mean	median	std.dev.
Math Score	0.261	0.011	0.096	0.087	0.056	0.375	0.025	0.170	0.175	0.075
Reading Score	0.471	0.029	0.302	0.324	0.113	0.230	0.007	0.106	0.086	0.055
Science Score	0.944	0.028	0.373	0.333	0.201	0.661	0.018	0.308	0.268	0.146
Fathers Educ	6.000	0.000	3.901	4.000	1.613	6.000	0.000	4.157	4.000	1.465
Mothers Educ	6.000	0.000	3.630	4.000	1.557	6.000	0.000	3.848	4.000	1.445
Family Type	4.000	1.000	1.978	2.000	0.521	3.000	1.000	1.854	2.000	0.372
Achievement	0.800	0.054	0.469	0.485	0.165	0.799	0.082	0.460	0.479	0.148
Circumstance	8.485	0.000	5.110	4.975	2.030	8.485	0.000	5.386	5.657	1.854

The raw data reveals improvements in parental circumstances over the period, though an increase in the prevalence of single parent families is evident, this is all reflected in the circumstance variable which shows increases in the mean and median and a reduction in the spread over the period. Coherent changes in the raw achievement variables are more difficult to discern but, as noted in the introduction, changes in national standards and testing methods were implemented in the intervening period. Basically the achievement variable to be used in this study shows a slight decline in the mean and median with a reduction in the spread.[9] To facilitate parsimony the circumstance variable was used to group the parental condition into three ordered categories: Lower, Middle and Upper. The circumstance class cutoffs were set at 4.5 and 6.6 for both observation years. Table 2 reports the class sizes that result from the categorization process.

Table 2: *Circumstance class sizes*

	2003	2009
Lower	0.430	0.367
Middle	0.343	0.391
Upper	0.227	0.242

Given identical class boundaries in both periods the circumstance distribution in 2009 stochastically dominates at the first order that of the 2003 period. Membership of the lower circumstance class has reduced significantly and membership of the Middle and Upper circumstance class has increased. The policy of elevating parental educational status seems to have worked. Turning to the determination of the achievement groups, the results of the various versions of the group number selection criteria are reported in Table 3. Visual representations of kernel[10] and semi-parametric versions of the distributions are provided in Figures 2 and 3.

Note that for all component comparisons GTR measures are significantly different at conventional levels of significance with the exception of the 4 *versus* 5 components 2003 comparison and all between year comparisons are significantly different. Both unweighted and importance weighted Transvariation measures yield the same conclusions when there is no penalization factor, 5 components in 2003 and 4 components in 2009. Similarly they yield the same conclusions under parsimony penalization, this time 3 components in 2003 and 4 components in 2009. Another way of viewing this

[9] Note that because the curriculum, examination and teaching methodologies had changed over the intervening period the 2003 and 2009 achievement scores are not really cardinally comparable.

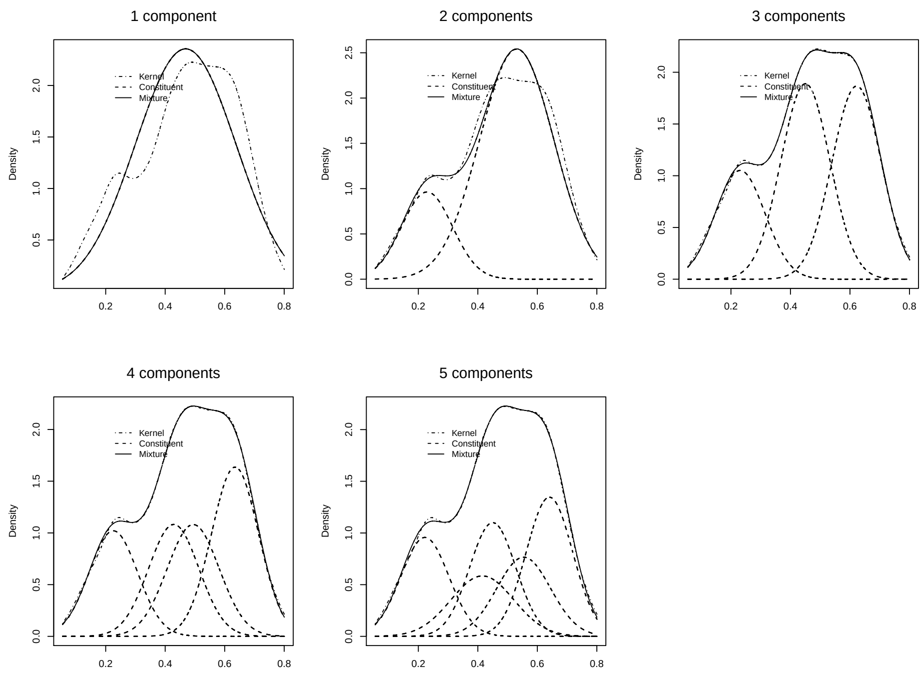
[10] A Gaussian kernel density estimator was employed. Silverman (1986) considers choice of the kernel a minor issue, what is crucial is bandwidth selection, since the estimated density is very sensitive to it. Following Jones, Marron and Sheather (1996), the bandwidth has been estimated using the plug-in procedure of Sheather and Jones (1991).

Table 3: *Kernel* vs *semi-parametric mixture transvariation measures 2003 and 2009*

No. classes	Year	GTR	GTR st.error	GTR + penalty	GTRIM	GTRIM + penalty
2	2003	0.113	0.005	0.127	0.097	0.113
	2009	0.072	0.002	0.079	0.063	0.073
3	2003	0.062	0.004	0.084	0.070	0.090
	2009	0.051	0.002	0.062	0.055	0.067
4	2003	0.057	0.004	0.086	0.064	0.094
	2009	0.040	0.002	0.055	0.047	0.065
5	2003	0.058	0.004	0.094	0.063	0.101
	2009	0.065	0.002	0.083	0.066	0.082

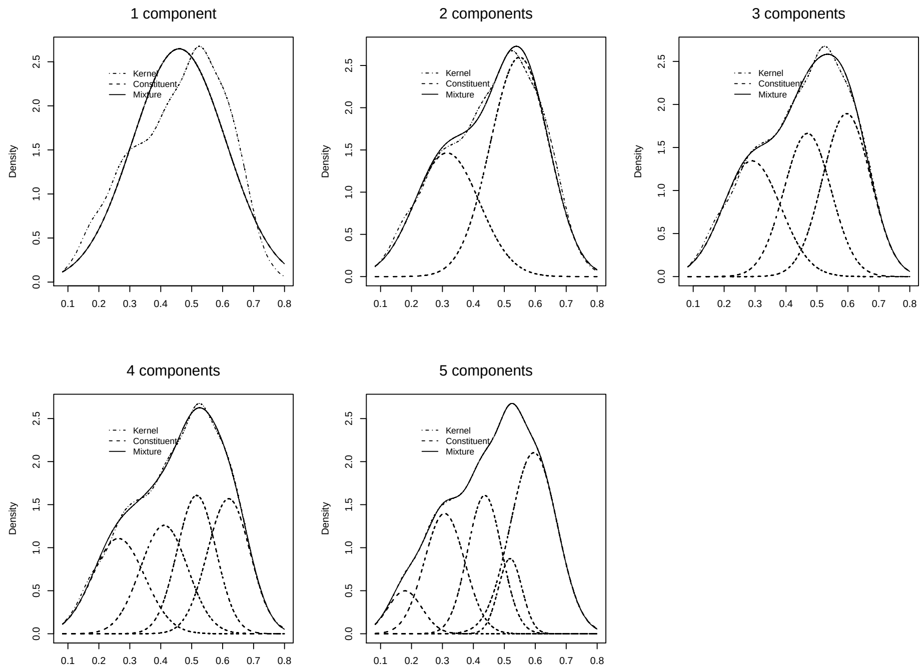
result is that parsimony penalization only has an impact on the choice of the number of components in 2003 (which was the smaller sample year). Thus, following the parsimony penalized criterion, the 3 achievement group model was selected for 2003 and the 4 achievement group model was selected for 2009.

Figure 2: *Kernel and mixture estimation of achievement in Germany: year 2003.*



The resultant achievement subgroup distributions for the two years are reported in Table 4. In 2003 all achievement groups have the same standard deviation suggesting

Figure 3: *Kernel and mixture estimation of achievement in Germany: year 2009.*



that the effort distribution is common to all groups. As may be observed the lowest achievement group has a similar population share in both 2003 and 2009 with a similar standard deviation (suggesting no change in the effort distribution) in both periods, the mean has however improved substantially from 0.231 to 0.263 (the standard normal statistic for the difference is 4.628). The Middle and Upper achievement groups of 2003 seem to have re-oriented themselves by 2009 into three equally sized groups identified as the Lower-Middle achievement group, the Upper-Middle achievement group and the High achievement group so that 2009 sees four roughly equal sized achievement groups. The effort distribution of the Lower-Middle achievement group has remained the same (an insignificant reduction in the standard deviation) whereas the effort distribution of the Upper-Middle and High achievement groups has tightened significantly in 2009.

For the purposes of establishing the presence of equality of opportunity, Lefranc *et* al. (2009) advocate examining the 2nd order stochastic dominance relationships of achievement distributions conditional on circumstance classes with the absence of dominance providing evidence of equality of opportunity. Table 5 presents evidence of 1st order dominance relationships of circumstance class conditional achievement

Table 4: *Achievement sub-group distributions: the components of the mixture models: years 2003 and 2009.*

	mean	std dev	weights
2003			
Low achievement group	0.231	0.075	0.220
Middle achievement group	0.450	0.074	0.391
High achievement group	0.622	0.074	0.389
2009			
Low achievement group	0.263	0.078	0.234
Lower-Middle achievement group	0.411	0.070	0.242
Upper-Middle achievement group	0.516	0.055	0.253
High achievement group	0.620	0.062	0.271

distributions for all groups in all years. That is to say achievement distributions of higher class circumstance groups always dominate those of lower class circumstance groups for all pairings in all years. Since first order dominance always prevails so will second order dominance, the equality of opportunity imperative has not been achieved in either year. However, using the area index (Anderson and Leo, 2015a), since the areas between the lower circumstance conditional achievement density and the high and middle circumstance conditional achievement densities have been reduced, there is some evidence of progress toward the equal opportunity goal at the lower end of the circumstance spectrum.[11]

Table 5: *Differences in achievement cumulative densities conditional on circumstance class: years 2003 and 2009.*

	2003			2009		
	$F(x\|C1)$ $-F(x\|C3)$	$F(x\|C1)$ $-F(x\|C2)$	$F(x\|C2)$ $-F(x\|C3)$	$F(x\|C1)$ $-F(x\|C3)$	$F(x\|C1)$ $-F(x\|C2)$	$F(x\|C2)$ $-F(x\|C3)$
Max	0.258	0.157	0.104	0.255	0.148	0.125
Min	0.000	0.000	0.000	0.000	0.000	0.000
Area	0.097	0.062	0.035	0.084	0.048	0.036

Turning to the circumstance to achievement transition matrices presented in Table 6 we see that the nature of the transition process has clearly changed over the period with the emergence of an additional achievement class. Furthermore, mobility has

[11]However, recall from footnote (7) that achievement scores are not cardinally comparable over the two observation periods so that technically 1st order dominance comparisons are all that are available to us. Non-the-less it is of interest to see what would result if they were comparable.

improved appreciably over the period: the TM mobility index moved from 0.804 in 2003 to 0.852 in 2009. The balance of probabilities convergence measure (PT) is equal to -0.220 for 2003 (a polarizing transition) and equal to 0.031 for 2009 (a convergent transition) favoring a move toward equality of opportunity over the observation period. The extent to which movement is upward as opposed to downward transiting (PUT) is 0.284 in 2003 as opposed to 0.106 in 2009, indicating that the move toward an equal opportunity goal has been at the expense of some upward mobility in the process.

Table 6: *Circumstance to achievement transitions: years 2003 and 2009.*

	2003 TM(M)=0.804		
	Circumstance 1	Circumstance 2	Circumstance 3
Low achievement group	0.320	0.169	0.108
	(16.528)	(7.802)	(4.071)
Middle achievement group	0.416	0.399	0.334
	(21.103)	(18.065)	(12.311)
High achievement group	0.265	0.432	0.558
	(12.697)	(18.496)	(19.449)
	2009 TM(M)=0.852		
	Circumstance 1	Circumstance 2	Circumstance 3
Low achievement group	0.338	0.191	0.144
	(23.960)	(14.018)	(8.274)
Lower-middle achievement group	0.269	0.256	0.180
	(25.806)	(25.424)	(14.072)
Upper-middle achievement group	0.174	0.269	0.419
	(12.401)	(19.827)	(24.245)
High achievement group	0.219	0.283	0.257
	(21.062)	(28.148)	(20.114)

Numbers in brackets indicate asymptotic t-values.

Finally the question arises as to whether the achievement distributions by circumstance class indicate any improvement over the 6 year period. Table 7 presents an analysis of the prevailing dominance relationships of the 2003 and 2009 achievement distributions for a given circumstance group. First and second order comparisons indicate no stochastic dominance at either order. However there is first order dominance of the 2009 over the 2003 distribution over the region 0–0.2, the lower tail of the achievement spectrum, which, following a lemma of Davidson and Duclos (2002)[12],

[12]Davidson and Duclos (2002) successfully demonstrated that, for distributions $f(x)$ and $g(x)$ de-

implies stochastic dominance at some higher order over the whole distribution. Essentially this indicates improvement in the lower tail of the achievement distribution of all parental circumstance classes, which implies the achievement distribution conditional on a circumstance class of 2009 stochastically dominates that of 2003 for all circumstance classes.

Table 7: *Stochastics dominance.*

$F_{i03}(x) - F_{i09}(x)$		Circumstance 1	Circumstance 2	Circumstance 3
First order	Maximums	0.044	0.019	0.007
comparison $i = 1$	Minimums	-0.048	-0.089	-0.072
Second order	Maximums	0.007	0.003	0.001
comparison $i = 2$	Minimums	-0.005	-0.019	-0.018
$0 - 0.2$ First order	Maximums	0.038	0.017	0.007
comparison	Minimums	0.000	0.000	0.000

5 Concluding Remarks

PISA assessment of academic attainments in Language, Mathematics and Science in Germany at the beginning of the 21st century shocked the nation, in short there was overwhelming evidence of a lack of equality of opportunity (OECD, 2011). This prompted extensive educational reforms at the federal and state levels over the ensuing years. Since a pure equality of opportunity objective is unlikely ever to be attained, the problem with assessing the effectiveness of these reforms is more one of measuring the progress toward the equality of opportunity imperative rather than determining whether or not transcendental optimality has been achieved. New tools for evaluating progress toward an equality of opportunity goal have been proposed and implemented in the context of the German reforms using PISA data for the years 2003 (immediately prior to the reforms being implemented) and 2009 (after the reforms had been implemented). Some progress in the imperative was detected especially at the lower end of the achievement spectrum which was the target of the reforms. Some structural change has taken place with the emergence of an additional achievement class over the

fined on a common interval $[a, b]$, if $f(x)$ stochastically dominated $g(x)$ at some order k over the interval $[a, c)$, $c < b$, then $f(x)$ would stochastically dominate $g(x)$ over $[a, b]$ at some order j where $j > k$.

period. Mobility was shown to have improved over the period and the gap between the achievements of the low-circumstance groups and their higher counterparts was shown to have narrowed. The nature of the transition process was fundamentally transformed with the emergence of an additional achievement class. It changed from a polarizing to a convergent process which can be seen to be equalizing the circumstances of future generations, though this was at the expense of some upward mobility.

References

Akaike, H. (1972). "Information theory and an extension of the maximum likelihood principle". *Proceedings of the 2nd International Symposium on Information Theory to Problems of Control and Information Theory*, 267–281.

Alkire, S. and Foster, J.E. (2011). "Understandings and misunderstandings of multidimensional poverty measurement", *Journal of Economic Inequality*, 9, 289–314.

Anderson, G. (2015). "Measuring Polarization and Convergence as Transitional Processes in the Absence of a Cardinal Ordering". *Working Paper 547*, University of Toronto, Department of Economics.

Anderson G. and Leo T.W. (2015). "Quantifying the Progress of Economic and Social Justice: Charting Changes in Equality of Opportunity in the United States 1960-2000". *Journal of Human Development and Capabilities*, forthcoming.

Anderson G. and Leo T.W. (2015a). "On Choosing the Second Best Outcome: Ranking Mutually Exclusive Alternatives When there is no Universally Dominant Solution". Mimeo, Department of Economics University of Toronto.

Anderson, G., Farcomeni, A., Pittau M.G., and Zelli R. (2015). "A new approach to measuring and studying the characteristics of class membership: The progress of poverty, inequality and polarization of income classes in urban China". *Journal of Econometrics*, Forthcoming.

Atkinson, A. B. (2012). "Public economics after the idea of justice". *Journal of Human Development and Capabilities*, 13(4), 521–536.

Atkinson, A.B., and Brandolini, A. (2013). "On the identification of the middle class", in *Income Inequality* (eds. J. C. Gornick and M. Jäntti), Stanford University Press, 77–100.

Banerjee, A.V., and Duflo, E. (2008). “What is middle class about the middle classes around the world?”, *Journal of Economic Perspectives*, 22(2), 3–28.

Beach, C.M., Richard P. Chaykowski, and George A. Slotsve (1997). “Inequality and Polarization of Male Earnings in the United States, 1968-1990”, *North American Journal of Economics and Finance*, 8(2), 135–151.

Brady, D.S., and Barber, H.A. (1948). “The Pattern of Food Expenditures”, *Review of Economics and Statistics*, 30, 198–206.

Citro, C.F. and Michael, R.T. (1995). *Measuring Poverty: A New Approach.* National Academy Press, Washington, DC.

Dardanoni, V. (1993). “On Measuring Social Mobility”, *Journal of Economic Theory*, 61,372–394.

Davidson R. and Duclos J.V. (2000). “Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality”, *Econometrica*, 68(6), 1435–1464.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). “Maximum likelihood from incomplete data via EM algorithm”, *Journal of the Royal Statistical Society*, B, 69, 1–38.

Easterly, W. (2001). “The Middle Class Consensus and Economic Development”, *Journal of Economic Growth*, 6(4), 317–335.

Fitoussi J-P, Sen, A. and Stiglitz, J.E. (2011). *Mismeasuring Our Lives: Why GDP Doesn’t Add Up*, The New Press, New York, London.

Foster, J.E. (1998). “Absolute versus Relative Poverty”, *American Economic Review*, 88(2), Papers and Proceedings, 335–341.

Gini, C. (1916). “Il concetto di transvariazione e le sue prime applicazioni”, *Giornale degli Economisti e Rivista di Statistica.* 1–55.

Grusky, D. and Kanbur, R. (eds.) (2006). *Poverty and Inequality.* Stanford University Press, Stanford, California.

Jones, M.C., Marron, J.S. and Sheather, S.J. (1996). “A brief survey of bandwidth selection for density estimation”, *Journal of the American Statistical Association*, 91, 401–407.

Kakwani, N. and Silber, J. (2008). *Quantitative Approaches to Multidimensional Poverty Measurement.* Palgrave Macmillan, London.

Kaufman, L. and Rousseeuw P.J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York.

Kobus, M, and Milos P. (2012). "Inequality decomposition by population subgroups for ordinal data", *Journal of Health Economics*, 31, 15–21.

Lefranc, A., Pistolesi, N., and Trannoy, A. (2009). "Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France". *Journal of Public Economics*, 93, 1189–1207.

Marron, J.S. and Wand, M.P. (1992). "Exact mean integrated squared error", *The Annals of Statistics*, 20, 712–736.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models.* Wiley, New York.

Milanovic, B. and Yitzhaki, S. (2002). "Decomposing World Income Distribution: Does The World Have A Middle Class?". *Review of Income and Wealth*, 48, 155–178.

Nussbaum, M.C. (2011). *Creating Capabilities: The Human Development Approach*, Harvard University Press.

OECD (2011). Chapter 9. Strong Performers and Successful Reformers in Education: Lessons from PISA for the United States. OECD Publishing. http://dx.doi.org/10.1787/978926409666 en.

Roemer, J. E. (1998). *Equality of Opportunity.* Cambridge, Massachusetts: Harvard University Press.

Roemer, J. E. (2006). *Democracy, Education and Equality: Graz Schumpeter Lectures.* Cambridge University Press. Econometric Society Monographs No. 40.

Roemer, J. E. (2010). "On responsibility-sensitive egalitarian allocation rules". Harvard University, Unpublished Manuscript.

Rossi. P.E. (2014). *Bayesian Non and Semi-parametric Methods and Applications.* Princeton University Press.

Saez, E. and Veall, M.R. (2005). "The Evolution of High Incomes in Northern America: Lessons from Canadian Evidence", *American Economic Review*, 95(3), 831–849.

Sen, A.K. (1983). "Poor Relatively Speaking", *Oxford Economic Papers*, 35, 153–169.

Sen, A.K. (2009). *The Idea of Justice.* Harvard University Press.

Sheather, S.J. and Jones, M.C. (1991). "A reliable data-based bandwidth selection method for kernel density estimation", *Journal of Royal Statistical Society (B)* 53, 683–690.

Shorrocks A. F., (1978). "The Measurement of Mobility" *Econometrica*, 46, 1013–1024.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Sutton, J. (1997). "Gibrat's Legacy", *Journal of Economic Literature*, 35, 40–59.

Townsend, (1985). "A Sociological Approach to the Measurement of Poverty - A Rejoinder to Professor Amartya Sen", *Oxford Economic Papers*, 37, 659–668.