**A Rational choice model of educational inequality**

Author(s): Breen, Richard, 1954-
Date 2001
Type Working Paper
Series Estudios = Working papers / Instituto Juan March de Estudios e Investigaciones, Centro de Estudios Avanzados en Ciencias Sociales 2001/166
City: Madrid
Publisher: Centro de Estudios Avanzados en Ciencias Sociales

**A RATIONAL CHOICE MODEL OF EDUCATIONAL INEQUALITY**

Richard Breen

Estudio/Working Paper 2001/166
October 2001

Richard Breen is Official Fellow at Nuffield College, Oxford University, and Fellow of the British Academy. This paper is based on a seminar he presented at the *Center for Advanced Study in the Social Sciences*, Juan March Institute, Madrid, on 24 April 2001, entitled "A Rational Choice Model of Educational Inequality".

## Abstract

Recent empirical research in the sociology of education has consistently reported three major findings. First, differentials in educational attainment between young people coming from different social classes have changed rather little, if at all, over the greater part of the twentieth century. Secondly, social class or other social origin effects seem to diminish in strength as students move to higher levels within the educational system. Thirdly, gender differences in educational attainment have declined dramatically over the past 25 years. In this paper I draw on the Breen and Goldthorpe (1997) model of educational decision making to provide explanations for the first two of these findings. I also show how the Breen and Goldthorpe model, as reformulated here, and drawing on Breen (1999), provides a sound behavioural model for the widely used Mare (1980; 1981) model of educational transitions.

Thanks to John Goldthorpe, Rob Mare, David McFarland and Juan Rafael Morillas for comments on earlier drafts. Earlier versions of the paper were presented at the meeting of ISA Research Committee 28 in Libourne, France, May 11-14th 2000; at a seminar at the Centro de Estudios Avanzados en Ciencias Sociales, Instituto Juan March, Madrid on April 24 2001; and at a seminar at UCLA on May 24 2001. Thanks to participants in all these events for comments and suggestions.

**A Rational Choice Model of Educational Inequality**

*Introduction*

The study of inequality and stratification is central to the discipline of sociology, and a great deal of sociological effort has gone into seeking to document and understand the processes by which patterns of inequality are reproduced or changed across generations - a topic sometimes termed 'social reproduction'. The educational system, as one of the central institutions of social reproduction, has accordingly been at the core of much of this work. Recent empirical research (for example the papers collected in Shavit and Blossfeld 1993) has consistently reported three major findings in this area. First, differentials in educational attainment between young people coming from different social classes have changed rather little, if at all, over the greater part of the twentieth century. This is not to ignore the increases in average levels of educational attainment over this period: although the mean level of attainment has indeed increased, differentials around this mean, according to class, have remained roughly constant. Secondly, social class, or other social origin, effects seem to diminish in strength as students move to higher levels within the educational system. For example, class differences in the rates of transition from lower- to higher-secondary education are usually greater than those found in the transition from, say, higher-secondary to tertiary education. Thirdly, gender differences in educational attainment have declined dramatically over the past 25 years.

In this paper I provide explanations for the first two of these findings. The paper draws on, and extends, Breen and Goldthorpe (1997). That paper presented a model of educational decision-making based on a rational choice approach with the aim of explaining why class inequalities in educational participation rates have remained largely unchanged. The central mechanism that Breen and Goldthorpe (henceforth B&G) used to account for this is 'relative risk aversion': that is, young people have, as their major educational goal, the acquisition of a level of education that will allow them to attain a class position at least as good as that of their family of origin. More simply, their chief concern is to avoid downward mobility. Following the publication of that paper a number of empirical studies have found support for the existence of this mechanism and for its role in accounting for differential educational attainment (Need and deJong 2001; Davies, Heinesen and Holm 1999). Another study (Schizzerotto 1997) has applied the model to explain cross-national differences in average levels of educational attainment, and the original paper has recently been reprinted in an edited collection on social stratification (Grusky 2001). Here I present a reformulation of the B&G model as a more general model of decision making when agents are faced with a sequence of risky choices and note that it is a special case of Kahneman and Tversky's (1979) prospect theory model.

The findings that the B&G model seeks to explain have all arisen from the application of a particular statistical model of educational transitions, developed by Robert Mare (1980; 1981). Recently the model has been criticized by Cameron and Heckman (1998). In response I devote the last part of the present paper to showing that the B&G approach provides a sound behavioural model for Mare's formulation – where, by behavioural model I mean a model of individual actions that, when aggregated, give rise to the observed empirical regularities.

The paper is in five sections. In the first I discuss research on educational inequality, and review both the methods used and the major findings that stand in need of explanation. The following two sections address the constancy of class differentials over time and declining social origin effects over the educational career. Section four turns to the use of the B&G model as a behavioural model underpinning the Mare model of educational transitions. There is a short concluding section. The contribution of this paper, over and above the original B&G paper, is threefold: first, it provides a simpler and more general formulation of the model; secondly, it explains the changing pattern of social origin effects over the educational career; and, lastly, drawing on the results in Breen (1999), it provides a behavioural foundation for the Mare model.

### *What needs to be explained*

Mare (1981: 73) points out that there are two aspects to educational stratification: one is the overall distribution of different levels of educational attainment, the other is the extent to which, given the overall distribution, there are differences between groups in their particular distributions over these levels. The latter issue requires that we pay attention to the relationship or association between education and social group membership. As Mare notes, these two aspects of educational stratification must be kept conceptually distinct: 'it is important, therefore, not to confound changes in the distribution of formal schooling with changes in the principles upon which schooling is allocated among groups' (1981: 74). Mare shows that previously used statistical models for the analysis of educational attainment failed to do this. For example, a simple linear probability model, in which the observed proportion who make a given educational transition is regressed on a set of social background variables, will tend to find that the latter have declining effects over time. This is because this measure conflates the true effect with the impact of aggregate educational expansion (Mare 1981: 75-6).

The Mare model distinguishes these two sorts of effect as follows. The educational career is conceptualized in terms of a progression over a set of sequential education levels.[1] In the American literature these are typically identified with school grades. At the completion of each level beyond compulsory schooling students either make the transition to the next level of education or they leave the educational system. Because each of the transitions that comprise the educational career is statistically independent, they can be modelled separately using a logit regression. This expresses the logarithm of the odds of making the transition from level k of the educational system to level k+1, conditional on having reached level k, as a linear function of J exogenous variables, $X_j$, whose values vary over individuals and, possibly, over transitions. As Mare (1993:353) has noted, such a model is a variant of the discrete time hazard rate model, where the hazard, in this case, is the probability of continuing to a further educational level, conditional on having reached a given level. In contrast to the usual hazard rate models, where each X variable has one coefficient, each variable may have a specific coefficient for each transition. This permits the taking into account of theoretically grounded hypotheses about differences in the effects of exogenous variables at different decision points, such as the idea that the influence of family background decreases with age. The model is preferable to one that focuses only on the highest

[1] There is a long-standing tradition in sociology of analysing educational careers as sequential transitions between grades or levels of education: for example, Boudon (1974).

educational level attained (for example, a multinomial or ordered probit model) because 'it corresponds better to the way that persons accumulate formal schooling, namely, in a sequence of irreversible steps' (Mare 1993:353).

Notwithstanding the many merits of the Mare model, and the fact that it has now become the standard approach to the analysis of educational inequality in quantitative sociology, there are, as Breen and Jonsson (2000) point out in a recent paper, some limitations to it. The most notable is the assumption that students progress through the educational system in a linear sequential mode, when, in fact, many school systems contain parallel branches of study. The choice is often not merely between staying and leaving but between staying in one of several types of education (such as academic or vocational) or leaving. Furthermore, while Mare's approach models the accumulation of education in a step-like sequence, it is by no means clear that the stepwise accumulation is a correct behavioural representation. This question has recently been raised by Cameron and Heckman (1998) and I discuss it more fully in the fourth section of this paper.

Most research into educational inequality nowadays uses the Mare model. The best example of its application is probably the volume edited by Shavit and Blossfeld (1993). This comprises studies of 13 countries: the USA, Germany, the Netherlands, Sweden, Britain, Italy, Switzerland, Taiwan, Japan, Czechoslovakia, Hungary, Poland and Israel. In all cases, retrospective data, collected mainly in the 1980s, were used to measure the extent of class and gender inequality in educational attainment in successive age cohorts born during the first two-thirds of the 20$^{th}$ century.[2]

The major findings of the Shavit and Blossfeld study were that, notwithstanding substantial expansion of educational systems during the century, particularly at the lower secondary level, in only two countries - Sweden and the Netherlands - was any reduction found in the strength of the associations, at the various transitions, between social class origins and educational attainment. The editors conclude (1993: 19) that, with the two

[2] Partial exceptions to this are Britain, where the data come from cohorts born between 1913 and 1952, Japan, where they come from cohorts born between 1905 and 1955, and Switzerland, where the data come from two cohorts one born in 1950, the other in 1960.

exceptions referred to, 'there has been *little change in socioeconomic inequality of educational opportunity'* (italics in original), and thus 'the impact of educational reforms on changes in educational stratification seems to be negligible' (1993: 21). In other words, although educational attainment levels have everywhere increased, the relative chances of continuing to further levels of the educational system have remained generally unchanged. They also found that '(w)ith the exception of Switzerland, the effects of social origins are strongest at the beginning of the educational career and then decline for subsequent educational transitions. In some countries (for example, the Netherlands, Sweden and Germany) the effects of social origin on the transitions to tertiary education are so small as to be insignificant' (Blossfeld and Shavit 1993: 18). Lastly, the constancy over time in class origin inequalities contrasts sharply with the position with respect to gender. Ten of the countries had data on both men and women and in all of them there was 'a substantial reduction' in male/ female differences in attainment.

The novelty of the Shavit and Blossfeld study lay in its attempt to make rigorous international comparisons by requiring a high level of consistency of approach from the authors of the country chapters. But its results echoed, to a large extent, earlier findings from the USA (Featherman and Hauser 1978), France (Garnier and Raffalovitch 1984), the Netherlands (Dronkers 1983), Britain (Halsey, Heath and Ridge 1980) and elsewhere (and later studies have arrived at very similar results).

*Constancy of class differentials via relative risk aversion*

The B&G model was developed in order to explain the first of these findings: the constancy of class inequalities over time. The model has three main elements:

1. The structure of the decision problem. B&G argue that within all educational systems there exist points at which young people (henceforth agents) have the choice of pursuing a more risky or a less risky option. The examples they give are the choice of an academic (risky) *versus* a vocational (less risky) track; and the choice of continuing to a

further educational level rather than leaving the educational system. Risk arises because of the pattern of expected utilities of the different choices and because there exists the possibility that agents who choose the more risky course may in fact fail to complete it.[3]

2. The existence of a threshold, $T_i$, that determines the $i^{th}$ agent's minimum acceptable level of educational attainment. All agents pursue a strategy of minimizing the probability of failing to reach $T_i$. B&G define $T_i$ to be a social class position at least as good as that from which the agent originated.

3. A set of subjective beliefs about the probability of succeeding in each of the risky options. B&G call their subjective belief parameter $\pi_i$ but in this paper it is referred to as $p_i$.

As pointed out in the introduction, the B&G model can in fact be given a simpler and more parsimonious expression which may open the possibility of adapting it to a wider class of dynamic choice problems. This is done as follows.

The educational system consists of a set of educational levels, k=1,…K, the first M of which are compulsory. At each level of post-compulsory education[4] there are two terminal educational outcomes (henceforth *outcomes*). These are failing ($F_k$ where k indicates the level of education); and leaving having succeeded and chosen not to continue to the next level, $L_k$. These terminal educational outcomes can be ranked in terms of (expected) utility and this ranking is agreed on by all agents. Specifically, $U(L_k) > U(L_{k-1})$ for all k (which means that higher levels of education, when successfully completed, have higher utility); and $U(F_k) \leq U(L_{k-1})$ (continuing to level k and failing is never preferred to leaving at level k-1).

Agents who fail at a given educational level must leave the educational system. Agents who complete a given educational level have the choice of leaving the system at that

---

[3] This may also be true of the less risky choice but for simplicity of exposition B&G assume that there is no chance of educational failure in this case.

[4] The discussion concerns only post-compulsory education.

point or continuing to the next level of education. Thus, the utility of succeeding at level k can be written:

$$U(S_k) = \max(U(L_k), V(k+1)) \qquad (1)$$

where $U(S_k)$ is the expected utility of succeeding at level k and V(k+1) is the expected utility of continuing to level k+1. V can be written:

$$V(k) = p_k U(S_k) + (1 - p_k) U(F_k) \qquad (2)$$

where $p_k$ is the agent's subjective probability of succeeding at level k[5].

The agent's decision rule is that she continues to level k of the educational system if

$$V(k) > U(L_{k-1}) \qquad (3)$$

If this inequality is not met[6] the agent leaves the educational system.[7] V(k) then depends on the returns to leaving the educational system at all higher levels and the subjective probabilities of succeeding at these higher levels. As B&G (1997: 289) note, this expected utility can be computed using backward induction.

*Definition*: Let ($r_1$, …, $r_M$) be a set of outcomes that are a (probabilistic) function of educational outcomes, $L_k$ and $F_k$, k=1,…,K; and $T_i$ be the outcome that acts as the threshold point for the i[th] agent.[8] For all outcomes with $U(r_m) \leq U(T)$, the change in utility, as an agent

[5] For convenience I drop the agent-specific subscript except where this would cause confusion.

[6] B&G (1997: 286) add a second condition. Agents will continue in education at level k only if both condition (3) and the condition $r_i > c_k$ are met, where r is the agent's family's resources and c is the cost of education at that level. Since the latter condition is rather straightforward I do not discuss it any further.

[7] Though I reiterate that stay and leave are used as illustrations of the generic choice between a more and a less ambitious option.

[8] I assume that this threshold is associated with a successful outcome, $L_k$.

moves to more preferred outcomes, is non-declining. For all outcomes having $U(r_m) > U(T)$ the marginal utility as an agent moves to higher ranked outcomes, is declining.[9]

On what basis do agents determine T? B&G assume that T is the securing of a class position equal to that of the agent's family of origin and so T is defined in terms of labour market outcomes which are themselves probabilistic functions of educational outcomes. In other cases T might simply be the highest level of education attained by one or other parent (as in Mare and Chang 1998 and Need and deJong 2001). The central point is that, given a definition of T, the fact that agents differ in where T is located implies that they will differ in their educational choices.

This model is a special case of 'prospect theory' (Kahneman and Tversky 1979), notably in the assumption that 'people generally perceive outcomes as gains and losses … defined relative to some neutral reference point' (274) which can be determined 'relative to an expectation or aspiration level' (286). Furthermore 'the value function for changes of wealth is normally concave above the reference point … and often convex below it' (278).

As an example of the application of the B&G model, I focus on the particular educational choice of staying in the system or leaving it as this is depicted in Figure 1 (reproduced from B&G 1997: 280). Here there are three possible educational outcomes - P (stay and succeed), F (stay and fail) and L (leave immediately) - and each is associated with known probabilities of gaining access to positions in three particular classes: the service class (denoted S*), working class (W*) and underclass (U*). These probabilities are denoted by $\alpha, \beta$ and $\gamma$. Parameters subscripted 1 refer to the conditional probability of entering the service class; 2 to the conditional probability of entering the working class; and the conditional probability of entering the underclass is equal to one minus the sum of the other two parameter values. The $\alpha$ parameters relate to transitions to class positions among those who continue and succeed at a given educational level; $\beta$ among those who fail; and $\gamma$ those who choose not to continue to that level. B&G assume that continuing and succeeding is a

[9] It may be the case that $T_i < M$ but this introduces no particular complications and so we ignore the possibility while recognizing that it may be empirically significant.

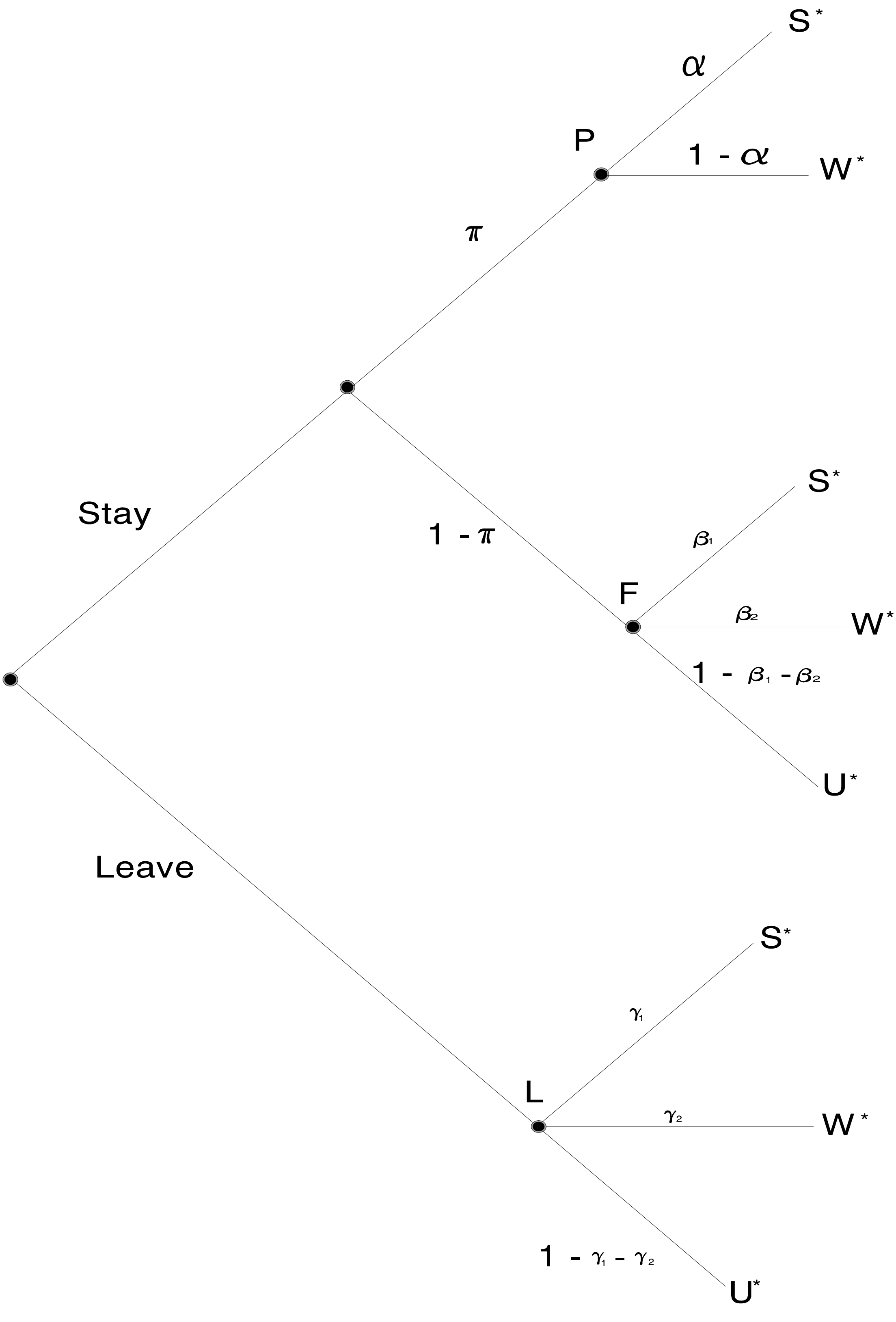
sure way of avoiding the underclass (i.e. $\alpha_1 + \alpha_2 = 1$) though this is not a necessary feature of the model. B&G (1997: 282) impose four assumptions about the relative sizes of these conditional probabilities. However, in the present reformulation of the model it emerges that only two of these are necessary.

(i) $\alpha > \beta_1$ and $\alpha > \gamma_1$. It is generally believed that remaining at school and succeeding affords a better chance of access to the service class than does remaining at school and failing or leaving school.

(ii) $\gamma_1 + \gamma_2 > \beta_1 + \beta_2$. Remaining at school and failing increases the chances of entering the underclass. This means that there is a risk involved in choosing to continue to the next level of education.

In comparing the probabilities of choosing to continue in education among agents of working class (W) and service class (S) origins I follow B&G in assuming that agents from both these classes display the same relative risk aversion in that their first priority is to minimize the risk of downward mobility. This implies that W agents seek to minimize the probability of arriving in U* while S agents seek to maximize the probability of ending in S*.

FIGURE 1



I now show that the definition of T, given above, together with the assumptions about the probabilistic relationship between educational outcomes and class positions, yields class differences in the probabilities of continuing to further levels of education.

Define, for S agents,

$$U(S^*) = U(W^*) + \lambda_1$$
$$U(W^*) = U(U^*) + \lambda_2$$

and for W agents

$$U(S^*) = U(W^*) + \varphi_1$$
$$U(W^*) = U(U^*) + \varphi_2$$

and set the thresholds at $T_S$=S* and $T_W$=W*. The λ and φ parameters thus capture the differences in utility between different class positions.

Let $p^*_S$ be the subjective probability below which an S agent will prefer the L option to continuing in education; and similarly for $p^*_W$ in the case of a W agent. The class difference that emerges from B&G's analysis is that $p^*_S < p^*_W$: that is, working class agents will require a higher subjective probability of success than will service class agents in order to choose to continue.

Write

$$p^*_S = \frac{(\gamma_1 - \beta_1)(U(U^*) + \lambda_1 + \lambda_2) + (\gamma_2 - \beta_2)(U(U^*) + \lambda_2) + (\beta_1 + \beta_2 - \gamma_1 - \gamma_2)U(U^*)}{(\alpha_1 - \beta_1)(U(U^*) + \lambda_1 + \lambda_2) + (\alpha_2 - \beta_2)(U(U^*) + \lambda_2) + (\beta_1 + \beta_2 - \alpha_1 - \alpha_2)U(U^*)}$$

In the case shown in Figure 1, $\alpha_1$ is equal to α and $\alpha_2$ is equal to B&G's 1-α.

A similar expression applies for $p^*_W$. Some algebra gives

$$p*_S < p*_W \Rightarrow \frac{\lambda_1(\gamma_1-\beta_1)+\lambda_2(\gamma_1+\gamma_2-\beta_1-\beta_2)}{\lambda_1(\alpha_1-\beta_1)+\lambda_2(\alpha_1+\alpha_2-\beta_1-\beta_2)} < \frac{\varphi_1(\gamma_1-\beta_1)+\varphi_2(\gamma_1+\gamma_2-\beta_1-\beta_2)}{\varphi_1(\alpha_1-\beta_1)+\varphi_2(\alpha_1+\alpha_2-\beta_1-\beta_2)}$$

and this reduces to

$$\begin{aligned}&\lambda_1\varphi_2[(\gamma_1-\beta_1)(\alpha_1+\alpha_2-\beta_1-\beta_2)-(\alpha_1-\beta_1)(\gamma_1+\gamma_2-\beta_1-\beta_2)] < \\ &\lambda_2\varphi_1[(\gamma_1-\beta_1)(\alpha_1+\alpha_2-\beta_1-\beta_2)-(\alpha_1-\beta_1)(\gamma_1+\gamma_2-\beta_1-\beta_2)]\end{aligned} \tag{4}$$

The assumption about the shape of the utility functions implies that

$$\lambda_1/\lambda_2 > \varphi_1/\varphi_2$$

In words: the change in utility between a W* and S* destination relative to the change in utility between U* and W* is larger for S agents (for whom these changes occur beneath their threshold) than for W agents. This means that the term in square brackets in (4) must be negative[10]: that is

$$(\gamma_1-\beta_1)(\alpha_1+\alpha_2-\beta_1-\beta_2) < (\alpha_1-\beta_1)(\gamma_1+\gamma_2-\beta_1-\beta_2) \tag{5}$$

---

[10] B&G examine two sets of values for the λ and ϕ parameters. Initially they set the utility of outcomes S* and W* to one for working class students while U(U*) is zero, and to U(S*)=1 for service class students and both U(W*) and U(U*) are set to zero. This implies $\lambda_1 = 1; \lambda_2 = 0; \varphi_1 = 0; \varphi_2 = 1$. Later B&G allow S agents to attach different utility to W* and U* and W agents to attach more utility to S* than to W*. So we have: U(S*)=1; U(W*)=0; and U(U*)=-x, for S agents; and U(S*)=x*; U(W*)=1; and U(U*)=0 for W agents. These give

$$\lambda_1 = 1; \lambda_2 = x$$
$$\varphi_1 = (x*-1); \varphi_2 = 1$$

The earlier formulation is consistent with the requirement that $\lambda_1/\lambda_2 > \varphi_1/\varphi_2$ and, if we set $1 \le x* < 2$ and $x \le 1$, so is the latter. Setting $x* = 1$ and $x = 0$ in the latter reduces to the earlier formulation.

By assumption (ii) above, $\beta_1 + \beta_2 < 1$ and so the second term on the left-hand side of (5) is positive. Assumption (i), $\alpha_1 > \beta_1$, ensures that the first term of the right hand side of (5) is positive and assumption (ii) ($\gamma_1 + \gamma_2 > \beta_1 + \beta_2$) means that the second term is also. Thus for the inequality in (5) to hold it must be the case that

$$\gamma_1(\alpha_2 - \beta_2) - \beta_1\alpha_2 < \gamma_2(\alpha_1 - \beta_1) - \alpha_1\beta_2$$

which reduces to

$$\gamma_1 < \beta_1 + \frac{(\alpha_1 - \beta_1)(\gamma_2 - \beta_2)}{(\alpha_2 - \beta_2)} \qquad (6)$$

B&G (1997: 282) say that their model imposes no assumptions about the relative magnitude of $\gamma_1$ and $\beta_1$. In fact the relationship between the two parameters itself depends upon the relationship between other pairs of parameters. The term $(\alpha_1 - \beta_1)$ is always positive, but $(\gamma_2 - \beta_2)$ and $(\alpha_2 - \beta_2)$ can be either negative or positive and thus $\gamma_1$ could be smaller or larger than $\beta_1$. In substantive terms, this means that it could, for example, be the case that a young person's chances of access to the service class are improved simply by acquiring more years of education, even if this does not lead to examination success. Alternatively - and, in many European educational systems, more plausibly - such time spent in education may be wasted in the sense that leaving school and embarking earlier on a career will yield a better chance of access to the service class.

The definition of a threshold, together with the assumptions about the $\alpha, \beta$ and $\gamma$ parameters, is sufficient to show that, given the strategies min (U*) for W agents and max (S*) for S agents, W and S agents have different probabilities of choosing the stay rather than leave option. Even controlling for differences in p (the subjective probability of succeeding if they continue in education) and in resources (with which to meet the costs of education), a higher proportion of W than S agents will prefer not to continue in education. This is because not continuing minimizes W agents' probability of downward mobility. Conversely, a large proportion of S agents will be obliged to continue because this maximizes their chances of acquiring a position in S*.

Class differences in the distribution of p and in the distribution of resources with which to meet the costs of education will then act to accentuate the class differences to which relative risk aversion gives rise. Such differences in p will exist because 'the mean level of ability is higher in the service class than in the working class' (B&G 1997: 285) and this difference is reflected in differences in educational performance which agents then use to form their expectations of success and failure in the future. 'If pupils' expectations about how well they will perform at the next level of education are upwardly bounded by how well they have performed in their most recent examination … then ability differences will be wholly captured in differences in the subjective parameter $\pi$' (B&G 1997: 286). The role of subjective expectations and their evolution over time has been developed in Breen (1999) and is discussed further below.

*Social origin inequalities throughout the educational career*

What are the implications of the relative risk assumption for patterns of class origin inequalities as agents progress through their educational careers? In particular, does relative risk aversion imply the pattern of declining effects that has been observed? To investigate this issue I begin with a simple example in which the educational career calls for two choices: whether to progress to level k=2 having completed level k=1; and whether to progress from level 2 to level 3. For agents of a given social class we can express the difference in the utilities of the educational outcomes as:

$$U(L_1) - U(F_2) = \lambda_1$$
$$U(F_3) - U(L_1) = \lambda_2$$
$$U(L_2) - U(F_3) = \lambda_3$$
$$U(L_3) - U(L_2) = \lambda_4$$

Following the earlier assumption all these λ parameters are non-negative.

To decide whether or not to proceed to level 2, an agent first computes whether or not she would continue to level 3, given that she succeeded at level 2. This yields a threshold value, $p^*_3$ given by $\lambda_3/(\lambda_3+\lambda_4)$. Define $E(p_{32})$ as the expected value of $p_3$ at the time when the agent is deciding whether or not to continue to level 2. Later I provide a model for the evolution of p that implies $E(p_{32}) = p_2$ and I assume this henceforth. If $E(p_{32}) > p^*_3$, the agent expects to continue to level 3 if she succeeds at level 2. It will be useful to express this in odds form: if $E(o_{32}) > o^*_3 \equiv \lambda_3/\lambda_4$ the agent would continue to level 3 if she succeeded at level 2.

The agent will continue to level 2 if

$$p_2 \max\{U(L_2), E(p_{32})U(S_3) + (1-E(p_{32}))U(F_3)\} + (1-p_2)U(F_2) > U(L_1)$$

This establishes a threshold

$$p^*_2 \equiv \frac{U(L_1)-U(F_2)}{\kappa_2 - U(F_2)} = \frac{\lambda_1}{\kappa_2 - U(F_2)}$$

where

$$\kappa_2 = \max\{U(L_2), E(p_{32})U(S_3) + (1-E(p_{32}))U(F_3)\}$$

And so

$$p^*_2 \le \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3}$$

From this result it follows that agents for whom $\kappa_2 = E(p_{32})U(S_3) + (1-E(p_{32}))U(F_3)$ will have a smaller value of $p^*_2$. For example, suppose there are agents of both the service and working class having thresholds $T_S=3$ and $T_W=2$. The choice of whether to continue to level 3, conditional on having succeeded at level 2, will be made by a higher proportion of S than of W agents, all else equal. This is because (as in the example of the previous section) non-declining marginal utility below the threshold means that $\lambda_3 / \lambda_4$ for the service class is less than the same quantity for working class agents. In certain circumstances, although the choice of whether or not to continue to level 2 from level 1 occurs beneath the threshold of both classes, a larger share of service class agents will choose to continue because, for more of the former, $\kappa_2$ will be equal to the larger value of $E(p_{32})U(S_3) + (1 - E(p_{32}))U(F_3)$ and so $p^*_2$ will be lower. Intuitively, a larger payoff later in the educational career makes it more attractive to remain in the system. But this will only occur when

$$\frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \geq \frac{\lambda_3}{\lambda_3 + \lambda_4}$$

That is to say, this will occur only if the threshold for progressing from level 2 to 3 is lower than the threshold for progressing from level 1 to 2 for agents who do not intend to continue beyond level 2. If these circumstances do not hold, and the later choice implies a higher threshold, the fact that those who expect to choose to progress to level 3 have a lower threshold at the earlier choice is of no significance since, by the assumption that $E(p_{32}) = p_2$, their value of $p_2$ would have exceeded the threshold in any case.

Now suppose that $T_W=1$. In this case both choices will occur above the working class threshold. In this case even among those agents for whom $\kappa_2 = U(L_2)$ a higher proportion from the service class will choose to continue (which would not be the case were $T_W=2$). This is because the threshold $p^*_2$ is set at a lower value by service class agents by virtue of their non-decreasing marginal utility. And, in the special circumstances explained in the preceding paragraph, the earlier argument also holds: a larger share of service class agents will have $\kappa_2 = E(p_{32})U(S_3) + (1 - E(p_{32}))U(F_3)$ and so agents of this class may have a higher rate of transition to level 2.

We can now state our first two results about social origin differences during the educational career:

(i) Holding constant the costs of education and the subjective probabilities of success, consider the difference between agents from groups having different thresholds in the probability of continuing from level k-1 to level k when this choice occurs below the threshold point for both these groups. In this case the probability of continuing will normally be the same for all agents, though there may be circumstances in which the transition rate is greater for agents with the higher threshold. Such a difference between agents with different thresholds would be larger, all else equal, for choices that were closer to the lower threshold.

(ii) Holding constant the costs of education and the subjective probabilities of success, the probability of continuing from level k-1 to level k of the educational system will be greater for agents for whom $k-1 < T$ than for those agents for whom $k-1 \geq T$. This follows immediately from the assumption of non-decreasing marginal utility below T and decreasing marginal utility above T. Class differences here (between thresholds) will generally be greater than class differences that occur below the thresholds (as in (i) above).

What happens when the choice occurs above the threshold of both classes? In that case we find that

(iii) Holding constant the costs of education and the subjective probabilities of success, consider the difference between agents from groups having different thresholds in the probability of continuing from level k-1 to level k when this choice occurs above the threshold point for all these groups and failure at k will not place an agent below her threshold. In this case the probability of continuing will be the same for all agents if absolute risk aversion is constant with increasing k, while, if absolute risk aversion declines with increasing k, agents from the group with the lower threshold will be more likely to make the transition to level k.

This follows because absolute risk aversion (ARA) is defined as the negative of the ratio of the second derivative of utility, with respect to educational level, to the first derivative: that is

$$ARA \equiv -\frac{U''(k)}{U'(k)}$$

If absolute risk aversion is declining then the ratio of the rate of change in the change in utility to the rate of change in utility increases. This implies that the ratio

$$\frac{U(S_k) - U(F_k)}{U(L_{k-1}) - U(F_k)}$$

is larger for a given k the further above T is k. Since the value of p that makes an agent indifferent between staying in education and leaving is

$$p^*_k \equiv \frac{U(L_{k-1}) - U(F_k)}{\kappa_k - U(F_k)} \qquad (7)$$

it follows that, all else equal, p* will be lower among agents for whom k is further above their threshold (where, as before,

$$\kappa_k = \max\{U(L_k), E(p_{k+1,k})U(S_{k+1}) + (1 - E(p_{k+1,k}))U(F_{k+1})\}$$

On the other hand, if absolute risk aversion is constant, then this ratio is constant with increasing k and so all agents for whom k occurs above their threshold will have the same p*.

Results (i) to (iii) imply nothing about the sequential pattern of conditional probabilities of making given educational choices - for example whether the probabilities for a given class will increase over educational levels. Rather they refer only to class *differences* at the same educational choice.

To clarify result (iii), consider a numerical example. Suppose that the returns to education (however we might measure them) are a simple linear function of years of education. Write utility as a function of years of education, Y, as follows:

$$U = -a/(y-T) \, for \; y > T$$

where a is a positive constant. This function yields decreasing ARA. Consider two agents, one with a threshold at y=10, the other at y=13, and they have the choice of leaving the educational system after 14 years of schooling or remaining for a further year at the end of which they either succeed or fail. If they succeed they have completed 15 years of education; if they fail they have completed 14 years' education. However, since they have incurred costs of remaining in school for an extra year, their lifetime returns to education are reduced by the equivalent of .25 of a year and so they leave with the equivalent of 13.75 years of education. In this case, using equation (9) with a=10, the agent with the threshold at y=10 requires a subjective probability of at least .25 in order to continue in education, while the agent with the higher threshold requires a p of at least .4.

Taking results (i) to (iii) we can now say something about class differences in educational choices throughout the educational career. The relative risk aversion implies, all else equal, that educational transition rates are unlikely to differ between agents of W and S in transitions that occur below both the thresholds, $T_W$ and $T_S$. To the extent that we do observe such differences these must be explained in terms that lie outside the reactive risk aversion assumption: for example in Boudon's 'primary effects' or in class differences in the subjective probabilty of success or in resource inequalities.

Class differences will be greatest for transitions that occur above $T_W$ and below $T_S$ since for S agents marginal utility is increasing in higher outcomes while for W agents it is declining. For a given decision, agents who are as yet some way beneath their threshold level of education will continue even given a rather low subjective probability of success at that level, $p_k$, because the expected gain in utility is nevertheless large. But agents for whom this decision occurs above their threshold - such as W agents - will require a much higher subjective probability of success in order to continue simply because the gain in utility is so much smaller.[11] Thus the model explains what Gambetta (1987: 171-2) describes as the contrast between the 'light-hearted' way in which middle-class families expose their children to the risk of educational failure and the extreme caution of the working class. The extreme case is given by the comparison between a class whose threshold is, say, the attainment of a University degree, and another whose threshold is reaching the minimum level of compulsory schooling. For all educational decisions that occur between these two points, the subjective probability of success required to continue will be very much greater for the latter than for the former.

The implication of relative risk aversion for transitions above the higher of the thresholds is that class inequalities will decline. Empirical analyses that report a declining impact of social origin at higher transitions are concerned with more than just two classes: indeed, social origin measures may also be continuous. While the exact pattern of aggregate or overall class inequalities will depend on the relative sizes of the classes, such inequalities will reach a maximum at a point between the lowest and highest class thresholds. Beyond this point class differences will decline, but this decline will be more marked when agents display declining absolute risk aversion than when they display constant absolute risk aversion.[12] In either case, however, the model presented here serves to explain the widely observed decrease in social origin effects as agents progress through the educational system.

11 And these different propensities will be reinforced if costs are a significant factor in the choice of whether or not to continue, given resource differences between S and W agents.

12 To see this consider agents from five classes, A to E, with $T_A < T_B < \ldots < T_E$. When class A passes its threshold its transition rate will be lower than that of all other classes (by (ii) above) while, among the other classes, rates will be the same (by (i) above). When class B passes its threshold, transition rates will be ordered B < A < C = D = E if absolute risk aversion is declining and A = B < C = D = E if absolute risk aversion is constant. This process continues until all classes have exceeded their threshold. In this example, given either constant or declining absolute risk aversion, aggregate inequality (measured as the rank correlation between transition rates and class, scored A=1 to E=5) reaches a peak once B has passed its threshold and then declines – more quickly if absolute risk aversion is declining.

*A behavioural basis for the Mare model*

Consider again equation (3), which is the criterion that determines whether or not an agent chooses to progress from level k-1 to level k of the educational system. If agents had perfect information about the returns to different levels of education, their costs, and their own probability of succeeding at different levels, they would know V(k) and $U(L_k)$ and would be able to determine, *ex ante*, their optimal level of educational attainment. As a result the model of sequential educational transitions introduced by Mare (1980, 1981) and widely used in analyses of educational attainment, would not be appropriate. Educational decisions would be determined prior to entry to the educational system and, given this, there would be no need to model each transition to a new grade or level separately. This is the core of the argument made in a recent paper by Cameron and Heckman (1998: 285-7). They suggest that agents pick that level of schooling that maximizes the difference between the costs and the discounted lifetime returns to schooling. They further assume that each agent's costs can be written as the product of an observed (by the economist) cost, multiplied by a person-specific and unobserved scalar random variable. This determines that the economist must use an ordered discrete-choice model to estimate agents' probabilities of reaching level k of the educational system. They then point out that Mare's statistical model is not compatible with their behavioural model.

In our notation, Cameron and Heckman suggest that agents maximize

$$U(L_k) - c_k$$

by choice of k, conditional on their resources and other background characteristics. The choice of schooling level is an investment decision made under certainty. By contrast, the model presented here assumes uncertainty in that $p_k$ is not known *ex ante*: rather it evolves as agents progresses through the educational system.

Breen (1999) has suggested an intuitively plausible way in which p might evolve. Agents have imperfect information about their probability of succeeding in school, where success can mean passing examinations, being commended by the teacher, and so on. They believe that they can influence this probability by exerting more or less effort, but they are unsure of how much difference this will make. At its simplest they can be thought of as entertaining two rival hypotheses about this: effort is very important in shaping the probability of success; and effort is not very important. These beliefs correspond to two possible states of the world which I call s' and s. Exerting effort is costly and agents apply Bayesian learning to the problem of learning which of the two states is true.

Formally the set-up is as follows. At time t, agents have an expected return from trying to succeed in education and exerting effort e(t), given by

$$EU(e(t)) = p(e(z_i(t), X_i), X_i)U(S) + (1 - p(e(z_i(t), X_i), X_i))U(F) - c(e(z_i(t), X_i)) \quad (8)$$

where U(S) and U(F) are the utilities of success and failure; e is the effort the $i^{th}$ agent exerts; c is the utility equivalent cost of exerting effort; and X is a vector of fixed individual level characteristics. p is the agent's subjective probability of succeeding and z captures beliefs about the returns to effort. It varies between zero and one and is the weight attached by an agent to the belief that s' is the true state of the world, and thus 1-z(t) is the weight attached to s being true. The amount of effort exerted, e, is a function of the belief in the returns to effort, z, and of individual level characteristics, X, and the latter can also shape the subjective probability directly. In equation (8) the expected return from trying to succeed is equal to the returns to success and failure, respectively, each weighted by their subjective probabilities of occurring, minus the utility equivalent cost of exerting effort, e.

Agents apply Bayesian updating to z. Given their belief in the returns to effort they choose how much effort to exert, they then observe the consequences and update their belief - that is, they change z, in the light of the observed outcomes of their actions:

Formally,

$$z(t+1) = \frac{pr(outcome \mid s', z(t))z(t)}{pr(outcome \mid s', z(t))z(t) + pr(outcome \mid s, z(t))(1 - z(t))}$$

The belief at t+1 (the posterior belief) depends on the prior belief (the belief at t) and the observed outcome. Agents assess the probability of the observed outcome having occurred given their belief and that s' is the true state of the world (this is pr (outcome | s', z(t))) and given that s is (pr (outcome | s, z(t))). If, for the given belief, the observed outcome is more likely under s than s', z(t+1) < z(t) and vice-versa. Since the prior belief, z(t) uses all the information available at t, the expected value of t in the next period is simply z(t). Thus the stochastic process $\{Z_{i,t}\}$ is a martingale and from this it follows that the process $\{P_{i,t}\}$ is also a martingale.

Suppose that agents gain utility from not trying to succeed in education, given by U(N). Equally this might be seen as a reservation level of utility. Agents decide whether, and how much, effort to exert first by solving equation (8) for the utility maximizing value of e (labeled $e^+$); substituting this into equation (8) and then choosing their course of action (or inaction) depending on whether $EU(e^+(t))$ (that is, the expected utility from trying to succeed, given effort $e^+$) is greater or less than U(N). But since X is a vector of fixed characteristics, $z^+$ can be substituted for $e^+$ and agents will only try to succeed in education if $z > z^+$. In other words, whether agents try or not depends (conditional on the various utilities) on their beliefs in the return to effort. But, if agents do not try, they do not learn. If the true state of the world is s' - that is, the return to effort is, in fact, high, - the Bayesian learning mechanism will give rise to a separation between those agents whose belief in the returns to effort becomes sufficiently low that they cease to update their beliefs and thus remain with a low belief in effort; and those agents whose beliefs remain above the threshold and converge towards (but not necessarily to) the true belief (which is z=1).[13] The simplest way in which this happens is when some agents with a low probability of succeeding prefer to terminate their education at a given level in preference to continuing.

Which of these groups an agent is found in depends on his or her initial beliefs, since the probability of the stochastic process $\{Z_t\}$ reaching any particular value depends on its initial value, z(0). If, for example, agents acquire their prior or initial belief from their parents, then if an agent's parents hold beliefs below the threshold $z^+$ the agent will never

[13] In the appendix there is an example of the Bayesian learning process.

learn. If his or her parents' beliefs are above but close to $z^+$ the agent has a high probability of coming to hold a belief $z(t) = z^+$; but for higher inherited beliefs this probability declines. In Breen (1999) it was pointed out that among adults class position may also depend upon effort and thus upon beliefs in the return to effort. The result will be that 'the working class, when compared with the middle class, will come to consist of a higher proportion of individuals who do not believe that effort is important' (Breen 1999: 470). If these beliefs are indeed passed from parents to children then this provides a plausible explanation for 'class differences in (average) beliefs that are often mistakenly described as class-specific norms or cultures' (Breen 1999: 470, parentheses added). These include such things as 'class differences in educational aspirations' (Coleman *et al* 1966; Jencks *et al* 1972); working-class culture of poverty, fatalism and inability to delay gratification (Hyman 1953; Lewis 1961, 1968; Pearlin 1971; Macleod 1995); working class reluctance to make sacrifices (Bourdieu 1974); middle-class culture in which parents place a high value on education and encourage their children accordingly (Halsey *et al* 1961), and similar (Breen 1999: 466). [14]

To integrate this Bayesian learning approach with the B&G model I assume that the process of updating z, and thus p, goes on throughout an agent's educational career. When agents decide whether or not to continue to level k of the system they evaluate the decision rule of equation (3) using their current value of p. The expected returns from continuing in education depend, *inter alia*, on the agent's belief in the probability of succeeding in further levels of education, but subjective probabilities of success evolve over time as a martingale. Thus, in computing V(k) agents work by backward induction, employing the known returns and costs to different educational outcomes and their current belief about p, their subjective probability of success. The B&G model implies that agents decide *ex-ante* on their necessary minimum level of educational attainment, but beyond this, they reevaluate whether or not to continue at each choice point in the educational system, in the light of the new information they have acquired since their last decision. In this case, the new information concerns their

[14] If leaving the educational system is the major reason for why agents cease to learn about the returns to effort within that system (and in the absence of other factors, such as cost barriers, that force agents to leave the system who otherwise would prefer to stay), then beliefs will be highly correlated with educational attainment. However, this correlation will be stronger within a social class than across the whole population of agents. This is because class differences in the utility attached to educational outcomes will, as shown by B&G, cause the threshold belief required to continue in education to differ between social classes. From this it follows that the belief in effort required to continue will also differ among them.

beliefs in the probability of succeeding, and because P is a stochastic process, its realizations cannot be known *ex ante*. Thus a statistical model of sequential decision making is appropriate.

## *Conclusion*

This paper has presented a reformulation of the B&G model of educational decision making and has used it to try to explain two widely reported empirical results in the analysis of educational inequality. The model also provides a behavioural foundation for the widely used Mare model of educational transitions. Mare himself noted that his model seems to correspond well to the way people accumulate schooling - namely in a series of steps. But the fact that this is so is simply a consequence of the structure of educational systems. It cannot discriminate between, for example, a situation in which agents decide *ex ante* how much education to accumulate and one in which they decide at the end of each level of education whether or not to continue to the next level. In the absence of an explicit behavioural model it is difficult to test specific hypotheses about the mechanisms underlying the various results to which the model's application has given rise. More generally, the aggregate observation that a statistical model is designed to capture (in this case sequential progression through the educational system) may be compatible with a number of different behavioural models each of which, may, in its turn, have different implications for policy. The behavioural foundation for the Mare model that the B&G model provides reflects the idea that agents' educational careers are the consequence both of *ex-ante* decisions by the families concerned (captured in the idea of an educational threshold) and of a process of 'learning by doing' involving beliefs about the probability of educational success.

APPENDIX *An illustration of Bayesian learning in education*

To illustrate the above argument, let the subjective probability p be defined as

$$\tilde{p} = (\tilde{g}X + \tilde{\theta}e) \qquad (1.1)$$

where g is the unknown weight given to (some known function of) the vector X and θ is the unknown return to effort. Assume that there are two possible states of the world, denoted s' and s. In s', effort is important, relative to the fixed factors, X, in determining educational success, while in s it is relatively unimportant. Thus we have

$$\theta' > \theta \text{ and } g > g'$$

Agents want to learn which of the states, s' or s, is the true state. Their beliefs about the returns to effort can be written

$$\tilde{\theta}(t) = z(t)\theta' + (1 - z(t))\theta \qquad (1.2a)$$

and about the returns to the fixed X factors as

$$\tilde{g}(t) = z(t)g' + (1 - z(t))g \qquad (1.2b)$$

where z is the weight agents attach to s' being the true state.

We suppose that when agents try to succeed in education, they exert effort and they succeed with probability $\tilde{p}$. Thus we can write the expected utility of exerting effort, e, as:

$$EU(e) = \tilde{p}(t)U(S) + (1 - \tilde{p}(t))U(F) - c(e) \qquad (1.3)$$

where U(S) and U(F) are the returns to success and failure and c(e) is the cost of exerting effort e. We also assume that agents gain utility from not exerting any effort (that is, not trying to succeed), equal to U(N) and that we have U(S) > U(N) > U(F).

Conditional on their belief, p(t), and a specific function for c(e), agents choose a level of effort to maximize (1.3) and they act only if $E(U,e^{+})$ – that is, the expected utility given the exertion of the maximizing level of effort, $e^{+}$ – exceeds U(N).

Consider the case in which

$$c(e,X)=\frac{e^2}{2cX} \qquad (1.4)$$

where c is a constant and the X variables are scored such that higher values reduce the costs of exerting effort. So, for example, one component of X might be ability and so higher ability reduces the costs of exerting effort. In this case the utility maximizing level of effort is

$$e^{+}(t)=\widetilde{\theta}(t)cX\Delta U$$

where $\Delta U \equiv U(S)-U(F)$.

From this it follows that there exists a threshold value of z, say $z^{+}$, beneath which agents will prefer to not act. In other words, if their belief in the returns to effort is sufficiently low, then they will not exert any effort. So, agents can be separated into two groups, depending on their value of z(t). If $z(t) > z^{+}$, agents will act and will update their beliefs using Bayes' rule. Their beliefs will eventually converge either to z=1 (putting full weight on the true state) or $z=z^{+}$. If $z(t) \leq z^{+}$ agents will not act and therefore will not update their beliefs: they will not learn.

**REFERENCES**

Blossfeld, H-P and Yossi Shavit. 1993. "Persisting Barriers: Changes in Educational Opportunities in Thirteen Countries" in Shavit, Yossi and Hans-Peter Blossfeld (eds.). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Col.: Westview Press, pages 1-24.

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

Breen, Richard. 1999. "Beliefs, Rational Choice and Bayesian Learning", *Rationality and Society,* 11, 4: 463-480.

Breen, Richard and John H, Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory", *Rationality and Society* 9: 3: 275-305; reprinted in Grusky, David B. (ed) 2001 *Social Stratification: Class, Race and Gender*, Boulder, Col: Westview Press, pages 459-70.

Breen, Richard and Jan O. Jonsson. 2000. 'A Multinomial Transition Model for Analyzing Educational Careers', *American Sociological Review* 65: 5: 754-72.

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, 2: 262-333.

Dronkers, J. 1983. 'Have Inequalities in Educational Opportunities Changed in the Netherlands? A Review of Empirical Evidence', *Netherlands Journal of Sociology*, 19: 133-50.

Davies, Richard, Eskil Heinesen and Anders Holm. 1999. "The relative risk aversion hypothesis of educational choice", Unpublished paper.

Featherman, D.L. and R.M. Hauser. 1978. *Opportunity and Change*, New York: Academic Press.

Gambetta, Diego. 1987. *Were They Pushed or Did They Jump? Individual Decision Mechanisms in Education*, Cambridge: Cambridge University Press.

Garnier, M.A. and L. Raffalovich. 1984. 'The Evolution of Equality of Educational Opportunity in France', *Sociology of Education*, 57: 1-11.

Goldthorpe, John H. 1996. "Class analysis and the reorientation of class theory: the case of persisting differentials in educational attainment", *British Journal of Sociology* 47, 3: 481-505.

Grusky, David B. (ed) 2001. *Social Stratification: Class, Race and Gender*, Boulder, Col: Westview Press.

Halsey, A.H., A.F Heath and J.M. Ridge. 1980. *Origins and Destinations: Family, Class and Education in Modern Britain*, Oxford: Clarendon Press.

Kahneman, Daniel and Amos Tversky. 1979. 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, 47, 2: 263-91.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions," *Journal of the American Statistical Association* 75, 1: 295-305.

Mare, Robert D. 1981. 'Change and Stability in Educational Stratification', *American Sociological Review*, 46,1: 72-87.

Mare, Robert D. 1993. 'Educational Stratification on Observed and Unobserved Components of Family Background' in Shavit, Yossi and Hans-Peter Blossfeld (eds) 1993, *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, Westview Press: Boulder, Co. pages 351-76.

Mare, Robert D. and Huey-Chi Chang. 1998. "Family Attainment Norms and Educational Attainment: New Models for School Transitions", Paper presented at RC28 meeting Taipei, Taiwan.

Need, Ariana and Uulkje De Jong. 2001. "Educational Differentials in the Netherlands: Testing Rational Action Theory", *Rationality and Society*, 13,1: 71-98.

Schizzerotto, Antonio. 1997. "Perche' in Italia ci sono pochi diplomati e pochi laureati? Vincoli strutturali e decisioni razionali degli attori come cause delle contenuta espansione della scolarite' superiore", *Polis* 11, 3: 345-65.

Shavit, Yossi and Hans-Peter Blossfeld (eds.). 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, Boulder, Col.: Westview Press.