The Role of Education in Intergenerational Social Mobility: Problems from Empirical Research in Sociology and some Theoretical Pointers from Economics

John H. Goldthorpe

WORKING PAPER 13-02
June 2013
Editor:

Erzsébet Bukodi

Department of Social Policy and Intervention
University of Oxford
Barnett House
32 Wellington Square
Oxford, OX1 2ER
Erzsebet.bukodi@spi.ox.ac.uk
The Role of Education in Intergenerational Social Mobility: Problems from Empirical Research in Sociology and some Theoretical Pointers from Economics

John H. Goldthorpe

Nuffield College and
Oxford Institute of Social Policy
Department of Social Policy and Intervention
University of Oxford

john.goldthorpe@nuffield.ox.ac.uk

Abstract

Insofar as research into intergenerational social mobility and its mediation via education has had any theoretical context, this has been provided by the functionalist theory of industrial and post-industrial society. However, the inadequacies of this theory have become increasingly apparent, and rather little appeal to it is now in fact made. Yet no alternative body of theory has emerged. I propose that in developing the micro-level theory that is needed, sociologists could, with advantage, adapt to their own purposes theories originating in the economics of labour markets and, in particular, three theories in some degree deviating from mainstream human capital theory: i.e. screening and signalling theory, job competition theory, and incentive-enhancing preference theory. In this way, explanations for findings going contrary to expectations under the functionalist theory can be advanced that can already be provided with some degree of empirical support and that would be readily open to further test.

Acknowledgements

This paper results from work undertaken on the project Education and Social Mobility across Cohorts (EMAC), with funding from the ESRC, Grant ES/IO38187/1. I am indebted to Erzsébet Bukodi (Principal Investigator) and to Walter Müller for helpful comments on previous versions.
Introduction

The role of individuals’ educational attainment in determining their chances of intergenerational social mobility has been the subject of extensive cross-national and technically sophisticated sociological research. With reference to the much discussed ‘OED triangle’, as shown in Figure 1, it would be generally agreed that E (educational attainment) is strongly associated with O (social origin), and that D (social destination) is strongly associated with E, but with a ‘direct’ association of O with D – i.e. one not mediated by E – still being present.\(^1\) However, it would be fair to say that, beyond this point, some non-negligible divergence exists in empirical findings, most notably concerning changing relations within the OED triangle over time. The situation is made more difficult by the fact that the development of relevant theory has not kept pace with that of research. Consequently, differing results – and their interpretations – tend to be simply counterposed, with little focused consideration being possible of their particular significance or wider implications.

![Figure 1. The OED triangle](image)

\(^1\) In the relevant literature, O, E and D are conceptualised and measured in differing ways: for example, O and D in terms of occupational prestige or ‘socio-economic’ status or of social class, E in terms of level of education attained or actual qualifications. However, I intend the argument of the paper to have general applicability.
Insofar as any theoretical context has existed for research into the pattern and evolution of relations within the OED triangle, it has been that provided by the functionalist theory of industrial, or post-industrial, societies (see esp. Treiman, 1970). From this standpoint, the ED association is seen as one that will steadily strengthen as a necessary response to the exigencies of the complex forms of technology and economic organisation that characterise such societies. If technological and economic dynamism are to be maintained, employing organisations must increasingly select personnel on the basis of the differing levels and kinds of knowledge and skill that educational qualifications serve to warrant. At the same time, while the theory would recognise that some degree of OE association is an inevitable consequence of family influences, the expectation would be that this association will tend to weaken. The rising demand for qualified personnel will require the expansion of educational systems and their progressive reform aimed at creating a greater equality of educational opportunity, so that all available human resources can be utilised as effectively as possible. In turn, then, the theory also leads to the expectation that as the ED association strengthens but the OE association weakens, the direct and also the overall OD associations will in turn weaken: i.e. ‘ascription’ will progressively give way to ‘achievement’ via education, and societies will become increasingly meritocratic and mobile.

However, a widely recognised, problem with such functionalist theory (e.g. Elster, 1979; Coleman, 1990) is that it is couched entirely at a macro-social level: that is, as indicated in the foregoing, in terms of societal exigencies and requirements and what is taken to follow from them. Nothing is said about how the postulated exigencies and requirements are actually recognised by social actors, nor about how this recognition leads to courses of action that produce responses of a functionally adequate kind. The theory can then only be directly tested at the macro-social level, and where empirical data collected at this level are not consistent with it, little indication can be gained of exactly why it is that the theory is failing.
Since the functionalist theory of industrialism or post-industrialism was first applied in the field of mobility research, empirical difficulties have in fact emerged in the following respects.

(1) At least in more recent periods, there is no indication of a tendency for the ED association to strengthen. In five out of six European societies studied over the late twentieth century (Breen ed., 2004; cf. Breen and Luijkx, 2004: 393) the ED association was found actually to \textit{weaken}: i.e. individuals’ social destinations – in this case defined in terms of class – were becoming \textit{less} closely related to their educational attainment\(^2\); and other research has produced findings on similar lines (Goldthorpe, 1996; Whelan and Layte, 2002; Goldthorpe and Jackson, 2008).

(2) It remains unclear whether a secular, cross-national tendency is present for the OE association to weaken. While recent research (Breen et al., 2009, 2010) has questioned earlier findings (Shavit and Blossfeld, eds., 1993) of an essential stability in this association and has indicated weakening tendencies in some societies, at least at lower educational levels, in other societies no consistent weakening trend is apparent (see e.g. Barone, 2009 – also for a useful review; Nordli Hansen, 2008; Bukodi and Goldthorpe, 2010; Saar, 2010). Further, questions have been raised about how far an apparent decline in the OE association may disappear if educational categories are appropriately differentiated in order to take account of the increasing, if perhaps informal, re-stratification of higher levels of educational systems that is

\(^2\) The five countries in question are Britain, France, Ireland, the Netherlands and Sweden. The ‘deviant’ case is Germany where the association was stable (for interesting discussion, see Klein, 2011). Sociologists’ findings that the class returns to education are tending, if anything, to decline sometimes causes surprise to economists in view of the argument, widely advanced in the economics literature, that ‘skill-biased technological change’ (SBTC) results in earnings returns steadily increasing. However, apart from the fact that movements in class and earnings returns need not be all that highly correlated, the SBTC argument seems unduly influenced by American experience, and a wider European perspective leads to far more qualified conclusions regarding earnings returns (see e.g. Harmon, Walker and Westergaard-Nielson, eds., 2001).
the typical accompaniment of their expansion (see e.g. Boliver, 2011; Ichou and Vallet, 2011).

(3) In the case of the OD association, whether considered in its direct – i.e. unmediated by E – form or overall, there is, as with the OE association, no decisive evidence of a general weakening trend. As regards the direct association, Bernardi (2012) provides a review of findings that differ by time or place in showing either some degree of weakening or no change. As regards the overall association, a weakening tendency has been documented for a number of societies (Breen ed., 2004; cf. Breen and Luijkx, 2004: 385-90) but in others (see e.g. Goldthorpe and Mills, 2008; Ishida and Miwa, 2008; Erola, 2009; Bukodi and Goldthorpe, 2010; Falcon, 2013) the overall OD association would appear to be best characterised as constant or as showing only ‘trendless fluctuation’ in the way earlier suggested by Erikson and Goldthorpe (1992).

(4) In addition to the three two-way associations found within the OED triangle, evidence is also regularly produced of a three-way association – i.e. of an interaction effect. Where it is believed that the OD association is weakening, an attractive interpretation of this effect suggests that this association is less strong the higher individuals’ level of educational attainment (e.g. Hout, 1988; Breen and Luijkx, 2004), so that the weakening in the OD association comes about as a compositional effect of the expansion of higher educational levels. However, if, as appears the case, it is a weakening in the ED association that is most regularly found, an alternative interpretation of the interaction effect, less congenial to the functionalist theory, would seem more relevant: i.e. that the ED association is less strong the more advantaged individuals’ origins (Goldthorpe and Jackson, 2008), so that the weakening in the ED association comes about as a compositional effect of occupational or class structural
change resulting in more individuals originating in more advantaged professional or managerial families.

Given research findings such as the foregoing that are to a significant degree inconsistent with the temporal and cross-national regularities that would be expected under the functionalist theory, students of social mobility have for the most part been content to leave them as brute empirical facts or, if some explanation is attempted, to appeal, in a rather ad hoc fashion, to ‘institutional variation’ – as, say, in the forms of welfare states or of educational and training systems. Any ‘universalising’ impulse, such as that which prompted the functionalist theory, thus disappears, and it is simply accepted that the role of education in social mobility, and its consequences, will be different in the differing institutional contexts that, across time and place, happen to prevail. However, before sociological explanation is thus abandoned in favour of explanation that is, ultimately, of a historical kind, an attempt would seem desirable at developing relevant theory at the micro-social level. That is, theory couched in terms of the action and interaction of individuals that aims to specify actual social processes, or mechanisms, through which the macro-level associations empirically demonstrable within the OED triangle might be generated, and including differences in the strength and direction of change of these associations. Theory of the kind in question could then, to the extent that it is empirically corroborated, help account for findings that are anomalous in relation to the functionalist theory, and at the same time lead to less ad hoc, more integrated accounts of exactly how institutional variation matters – i.e. as providing the differing conditions under which the theory operates (cf. Müller and Gangl, 2003).

In the present paper, I seek to make a contribution in this regard by examining what might be gained by way of filling the theoretical void in which social mobility researchers presently operate from a closer, if critical, engagement with the work of economists, whose disciplinary
tendency – for better or worse – has been to move from theory, and especially micro-level theory, to empirical research rather than *vice versa*, and also to express such theory in terms of action that is, in some sense, rational (cf. Coleman, 1986). The theories on which I shall focus are ones that were primarily developed in order to explain the economists’ equivalent of the ED association – i.e. the earnings returns to education. But as in recent years economists have become increasingly interested in intergenerational earnings or income mobility (OD), their theories have been correspondingly extended to deal also with the association between parental earnings or income and children’s educational attainment (OE).

The plan of the paper is straightforward. I begin with a brief discussion of what stands as the mainstream economic theory in the area in question, namely, human capital theory. For reasons I indicate, this is unlikely to have any large appeal for sociologists. But I then go on to consider three further theories that have been developed in order either to complement or in certain respects to challenge human capital theory, and that, for present purposes, would appear to have more of interest to offer to sociologists. These are: screening and signalling theory, job competition theory, and incentive-enhancing preference theory. I aim to show how each of these theories can, potentially at least, contribute something of value to the understanding of relations within the OED triangle. A good deal of conjecture is obviously involved; but I am at the same time generally able to point to empirical findings that lend a degree of support for the ideas put forward and, more importantly, to indicate lines of research that would contribute to their further testing.
Human capital theory

Human capital theory as it developed from the late 1950s (Mincer, 1958; Schultz, 1962; Becker, 1964) can be seen as part of a continuing attempt to integrate the treatment of labour, as a factor of production, within the general framework of neoclassical economics. In earlier analyses the supply of labour was in effect envisaged as a stream of infinitely divisible homogenous units – rather than as coming, as Granovetter has aptly put it (1981, p. 17), ‘in inconvenient lumps called workers’. While from this position it could be shown that (under certain assumptions) the market wage would equal the marginal product of labour, little could be said theoretically about wage differences among workers – other than that some were apparently able to supply more units of the ‘common stuff’ of labour than were others (cf. Phelps Brown, 1977: ch. 1). Introducing the idea of human capital thus marked a major advance in providing a way of recognising both the existence of workers and the heterogeneity of their labour. Under the theory, individuals do not simply take their labour power as given but, as rational actors, seek to ‘invest in themselves’ so as to increase their productive capacities, and thus their potential earnings, up to the limits set by their natural endowments. And they do this following a similar calculus to that used in investment in physical capital. For example, in the case of investment in education as a means of enhancing human capital, individuals will make that investment, in terms of the direct costs incurred and of earnings forgone while in education, that will maximise their – appropriately discounted – expected lifetime earnings.

Extending the theory to cover the part played by education in intergenerational earnings mobility – or immobility – is then relatively straightforward (Becker and Tomes, 1986). Parents, being altruistic towards their children, will help them invest in their human capital via education, over and above any public provision, by foregoing some of their own
consumption. Children of higher income parents will benefit more in this way than will children of lower income parents since more resources will be available for them and also since they are likely, for both genetic and sociocultural reasons, to have greater academic potential. With smoothly functioning capital markets, inequalities in educational opportunities resulting from differences in parental incomes can be offset by borrowing to finance education. Parents can ‘equate the market interest rate on borrowing with the present value of the marginal return to investing in offspring’ (Blanden et al., 2011, p. 30). Conversely, though, inequalities will be accentuated insofar as imperfections in capital markets result in credit constraints that limit the possibilities for poorer children to have adequate investment in their education.

Sociologists have well-documented reservations about human capital theory. Some of these concern its assumptions of hyper-rationality in educational choice - the assumptions in effect of adolescent or parental ‘econometricians’ (Manski, 1993; cf. Breen and Goldthorpe, 1997). However, for present purposes, what is most seriously questioned is whether labour can be usefully conceptualised as just another factor of production. Labour is physically inseparable from the individual men and women who provide it, and the latter do not – the self-employed apart – sell their labour on the market in any direct sense. Rather, workers are offered and take a job, and in so doing enter into a contract with an employer. The general form of the contract provides that, in return for remuneration, workers will accept the authority of the employer in determining the actual work-tasks they should perform, the methods and procedures they should follow, the results they should seek to achieve etc. Thus, employment contracts are not so much concerned with the sale of a commodity as with the regulation of the social relationships involved in defining and performing jobs.
In the one case in which human capital theory does appear to have exerted some influence on the sociological study of mobility – that of ‘status attainment research’ (e.g. Blau and Duncan, 1967; Featherman and Hauser, 1978) – criticism came in fact to centre on the neglect of the part played by job structures in conditioning labour market outcomes. Following the supply-side emphasis of human capital theory, this research concentrated on the effects of individuals’ education on their ‘socioeconomic’ status attainment, as indexed by occupation – but under the implicit assumption, as Granovetter (1981, p. 15) observed, that job structures would be ‘infinitely malleable’ in response to changes in the distribution of education. A dissatisfaction with this assumption was then one of the main sources of the shift that subsequently occurred in the way in which the context of social mobility was conceptualised by sociologists: that is, a shift away from the idea of an essentially amorphous ‘socioeconomic’ status hierarchy in favour of that of historically and cross-nationally specific occupational or wider social class structures, recognised as being to an important extent determined exogenously to characteristics of labour supply.

However, if human capital theory has given rise to largely negative reactions from sociologists, attempts made by economists, of a more or less critical kind, to go beyond it could reward a more positive sociological reception.

**Screening and signalling theory**

Screening and signalling theory could be regarded as the most influential response to human capital theory that has so far developed within modern labour economics. From a sociological standpoint, what is of immediate interest about such theory is that it brings employers as well as workers into the analysis and recognises, at least in part, the problematic nature of the employment contract.
For exponents of screening theory (e.g. Arrow, 1973; Stiglitz, 1975), a major difficulty for employers, and especially when dealing with new entrants to the labour market, is that they have to engage in employment contracts on the basis of only very limited information about individuals’ productive capacities. The suggestion then is that what leads employers to rely on educational qualifications in their selection procedures is not, or at least not only, the belief that such qualifications certify that their holders have acquired relevant knowledge and skills. Employers can also take qualifications as a source of ‘statistical’ information about certain more basic attributes of potential employees. That is to say, they can use educational attainment as a means of screening, or sorting, individuals whom they might hire in regard to a range of personal characteristics that are not readily observable before the employment contract but that will be important for worker productivity within the contract. In particular, an individual’s educational record can be informative about his or her ability and readiness to acquire knowledge and skills (‘trainability’), to sustain effort, and to act in a disciplined and co-operative way. In this perspective, therefore, the significance of education for employers is not that it actually creates the attributes that they value in employees, as human capital theory would imply, but that an individual’s educational attainment helps employers to identify such attributes.

Correspondingly, then, from the side of workers, education can be regarded not so much as an investment in human capital per se as in the means through which they can signal to employers that they do have productive potential that, prior to hiring, they cannot directly demonstrate (Spence, 1973, 1981). However, it is important to note here that for such signalling to function reliably – i.e. not to be open to spurious use – it must be the case that its costs are less for individuals who do in fact show a greater productive capacity already in

---

3 I follow Weiss (1995) in referring to screening where (uninformed) employers are taken as first movers and to signalling where (informed) potential employees are taken as first movers.
their education – for example, who learn more easily – than for those showing a lower capacity; otherwise, all would invest in the same signal.

This last point matters because it seems to have been often supposed by sociologists that screening and signalling theory breaks the link, central to human capital theory, between education and productivity, and can thus be aligned with theories of ‘credentialism’ in which education essentially serves group or class, rather than organisational or societal, interests via processes of ‘social closure’ or ‘social reproduction’ (e.g. Collins, 1979; Bourdieu, 1984). But this is not the case. The link with productivity is central to both theories, as also is the further link with earnings. What is at issue is not whether, but how, education relates to productivity. Are an individual’s educational qualifications to be understood as warrants of an actually increased productive capacity or as signals of productive potential that employers could realise?

Economists have concentrated on devising empirical analyses that could evaluate the rival claims of screening and signalling and of human capital theory in accounting for earnings returns to education (see e.g. Chevalier et al., 2004). However, for sociologists, the attraction of screening and signalling theory is that it suggests explanations for already established findings on relations within the OED triangle which, so far, have been left unilluminated by theory, and explanations that would, moreover, appear for the most part to be readily open to empirical test.

To begin with, the finding in most obvious contradiction with functionalist theory, that in most advanced societies today the ED association appears to weaken rather than to strengthen, could be the result of educational qualifications increasingly playing a screening and signalling rather than a certifying role. Insofar as education is regarded as producing human capital – i.e. as in itself enhancing individuals’ productivity – it should continuously favour
career progression, and should help offset, where necessary, any relatively low-level entry into the labour market (Sicherman and Galor, 1990). But insofar as education is regarded more as a means of screening for productive potential, its effects should be concentrated at the time of labour market entry – i.e. on individuals’ chances of ‘getting in the door’ (Bills, 1988) – with the degree of their subsequent advancement then depending primarily on how far the potential that their education served, more or less reliably, to signal is actually expressed (Lange, 2007). In other words, while in both cases an association will exist between individuals’ qualifications and the positions – e.g. the social class or status positions – that they eventually achieve, where education serves as the basis of screening and signalling this association, while relatively strong in early working life, could be expected to become subsequently looser.

Further, proponents of screening and signalling theory have from the start emphasised that its applicability may in fact vary by economic sector and type of work (Spence, 1973: 349, 358-9; Riley, 2001: 460-3). For example, in the case of professionals, technicians or workers in skilled trades, it could be supposed that employers are primarily concerned that the individuals they employ do possess a particular range of knowledge and skills, and see appropriate qualifications as adequately certifying that such human capital has been acquired.\(^4\) In contrast, with non-technical managerial positions or with various ‘people processing’ occupations in, say, sales or personal services, where what constitutes relevant knowledge and skills is less easily defined and likely to be more firm-specific, employers may be more concerned with non-observable characteristics for which educational attainment could help them screen (Jackson, 2002; Jackson, Goldthorpe and Mills, 2005; Klein, 2011).

\(^4\) In some cases – most typically with professions – particular formal qualifications may of course be legally required.
If, therefore, types of work in which the screening – and in turn the signalling – role of education tends to predominate are in expansion, a general weakening of the ED association could result simply as a compositional effect. And it is of course well-established that in advanced economies employment growth does primarily occur in services sectors, while further evidence indicates that over recent decades this growth has been more marked in sales and personal services than in services within the ‘knowledge economy’ (Esping-Andersen et al., 1993; cf. Chang, 2010: ch. 17).

The force of an explanation on these lines for the tendency for the ED association to weaken could be more directly tested on the basis of work-history data. One could investigate how far differences of the kind expected by sector and type of work do show up in the degree to which the course of individuals’ worklife mobility, leading to their destination positions, is consistently predictable from their qualifications at entry into employment (human capital theory) or, rather, becomes progressively less so (screening and signalling theory).

Moreover, if education can serve employers as a means of identifying productive capacity – as distinct from actually creating it – there seems little reason why other individual attributes should not operate in a similar way. Thus, while educational attainment may provide a good screening criterion as regards general desiderata such as trainability, capacity for effort etc., with certain types of work other, more specific attributes may also be relevant for productivity. In particular, in more high-value sales and personal services, employers may regard it as important that employees have some familiarity and indeed empathy with the mores and lifestyles of the social groups from which customers or clients are chiefly drawn. Thus, aspects of job applicants’ own social backgrounds, activities and networks, as reported in their CVs or revealed in interview, may become the basis of screening, on the one hand, and of signalling, on the other. In this way too, then, the ED association could be weakened.
And, in addition, in the light of this possibility, explanations may also be suggested for two other regularities in relations within the OED triangle that the functionalist theory cannot accommodate.

First, if the interaction effect that regularly shows up is interpreted as meaning that the ED association is weaker for children of more advantaged than for those of less advantaged origins, then this could be seen to result from the former benefiting more from the signals of their social background that they can provide, and that can compensate, where necessary, for only modest educational attainments. There is evidence already available, for Britain at least, that employers are responsive to the social backgrounds of job seekers (Jackson, 2009), and, further, that children of parents in the professional and managerial salariat who have obtained only lower-level qualifications do in fact have clearly better chances of accessing the salariat themselves than working-class children with similar qualifications (Goldthorpe and Jackson, 2008). Moreover, this same study also provides indications – on which further, more focused, research could readily check – that it is disproportionately via careers made in the high-value sales and personal services sectors, leading to managerial positions, that this access to the salariat is gained.

Second, if it is in fact the case that for entry into some expanding range of occupations screening and signalling occurs in regard to individual attributes that are yet more closely linked to social origins than is educational attainment, this could be a factor helping to explain why a sustained weakening in the direct OD association, and in turn, perhaps, in the overall OD association, is not apparent across all modern societies alike: i.e. the operation of the labour market may still in certain respects favour ascription over achievement.
Job competition theory

Job competition theory is largely attributable to Thurow (1976; 1983: ch. 7). If human capital theory represents an attempt to more fully integrate the treatment of labour as a factor of production within the general theoretical framework of neoclassical economics, Thurow’s work can be seen as an explicit rejection of this attempt, and on grounds that most sociologists would find congenial (see esp. 1983: 173-81). Labour, he argues, is not bought and sold on the market but rather ‘rented’ by workers to employers in the context of a job and an employment contract. Further, Thurow argues, wages, as the price of labour, do not attach to workers themselves but rather to jobs, and if the market wage equals the marginal product of labour, then the marginal product resides in the job, not in the individual.

Underlying this position is Thurow’s belief that labour markets operate in a way far removed from the model set up by human capital theorists. This model supposes that workers acquire knowledge and skills ‘exogenously’ through formal education and training – i.e. invest in their human capital – and then bring their knowledge and skills into the labour market where they compete with each other on the basis of the wages they are willing to accept. But, as against this, Thurow claims that most workers acquire most of the knowledge and skills that they use in their jobs only after they obtain them: that is, through on-the-job training of a formal or informal kind. Thus, the labour market should be understood not, or at least not only, as a market in which workers with differing kinds and levels of human capital compete with each other over wages but rather as a ‘training market’ in which workers compete with each other for jobs, understood as ‘training opportunities’ that offer differing lifetime earnings prospects. The allocation of workers to jobs can then be envisaged as taking place via the matching of two ‘queues’: on the one hand, the ‘job queue’ or the ordering of available jobs by employers according to the level of their requirements and rewards; and, on the other hand,
the ‘labour queue’ or the ordering of workers according to their attractiveness to employers in terms of the likely training costs involved in employing them at a given level in the job queue.

As regards the way in which this matching is actually carried out, Thurow in effect takes over a version of screening theory. Employers will grade workers within the labour queue in the light of a range of ‘background characteristics’ that are taken as good indicators of the training costs that these workers will impose in relation to different jobs. Since educational attainment can be regarded as reflecting both trainability and discipline, it will tend to be the indicator that is given most weight for most types of job, although in addition such characteristics as age, sex and lifestyle characteristics may all play a part. Employers can then be thought of as starting at the top of their job queue and offering employment first to those workers in the labour queue who are indicated as having the lowest training costs for high-level jobs. If there is a shortage of such workers, employers will be forced to move down to workers of a lower – i.e. more costly – grade in the labour queue; or, if there is a surplus of higher-grade workers, employers will be able to engage some of these in lower-level jobs, and so on. Workers will respond by taking the best jobs available to them: that is, those offering the best training opportunities in terms of lifetime earnings prospects, although, perhaps, with non-monetary rewards also being taken into account.

In the same way as screening and signalling theory, Thurow’s job competition theory would seem to offer a more compelling challenge to human capital theory with certain types of work than with others, and in particular as regards his basic claim that the knowledge and skills that workers use are essentially acquired through on-the-job training. It would, for example, appear somewhat implausible to suppose that in the case of employees in professional, technical or craft occupations, their prior education and training do little more than provide an indicator of their likely responsiveness to further training following their entry into work.
Nonetheless, there are at least two features of job competition theory that appear of potential value to sociologists in search of a better understanding of the micro-level social processes that underlie the macro-level relationships demonstrable within the OED triangle.

First, job competition theory gives more adequate attention than either human capital or screening and signalling theory to the relation between demand and supply in labour markets. Thurow stresses that job, and thus training, opportunities are created only when there is a demand for the kind of work involved. He recognises that, to some extent, supply may create its own demand: i.e. employers may make technological and organisational adaptations, and thus change the nature of the job queue, in response to changes in the quality of the labour queue. But, he insists, a wide range of other factors affect demand so that, while the two queues do not move entirely independently, neither will they be ‘mirror images of each other’ (1983, p. 97). Moreover, on the supply side, educational expansion may be driven not by a perceived economic need for an increase in better qualified personnel but simply by rising expectations from intending students.

It is therefore possible for problems not only of under-qualification but also of over-qualification to arise. In the latter case, some individuals with higher-level qualifications – say, degrees – end up in lower positions than previously in the labour queue and, in turn, ‘bump down’ further those with lower-level qualifications. Under human capital theory, such problems will represent no more than temporary disequilibria, since employers will always change their production processes in order to exploit unused human capital. However, under job competition theory, over-qualification can be a persisting phenomenon, both at the aggregate level and over the course of individuals’ employment. And evidence has in fact of late accumulated of its persistence in both of these respects – evidence that has led to a much
greater interest in job competition theory than it previously attracted (for reviews see McGuinness, 2006; Quintini, 2011).

From a sociological standpoint, what is in this way underlined is Granovetter’s point, previously referred to, that the structure of jobs is not ‘infinitely malleable’ to educational supply but may, rather, condition the returns that education brings in the labour market. And, for present purposes, what is of chief interest is that another possible explanation of the weakening ED association is thus indicated. Given an excess supply of higher level qualifications, the returns they bring in terms of D – i.e. the class or status positions that individuals eventually achieve – will tend, on average, to decline as employers not only make more refined distinctions within categories of qualification (Green and Zhu, 2010) but also discriminate by drawing more on other kinds of information. There is, for example, already evidence of employers making increasing use of the results of psychometric tests of both cognitive ability and personality (Jenkins and Wolf; 2005), while, as already suggested, social background characteristics may, for certain types of work, take on renewed importance (cf. also Brown and Hesketh, 2004). So far then as future research is concerned, what is here pointed to is the need for studies of social mobility to be extended so as to incorporate the recruitment and promotion policies of employers in a far more systematic way than has been the case hitherto (cf. Bills, 2003).

The second way in which job competition theory offers the possibility of a better understanding of the role of education in social mobility derives from the central idea of the matching of job and labour queues. It is an important implication of this idea that educational attainment should be viewed in relative rather than absolute terms or, that is, as a positional good (Hirsch, 1977; Wolf, 2002: ch. 8). Insofar as education serves as a key ‘background characteristic’ for employers in selecting employees, the level of job in the job queue that is
available to an individual will depend not simply on how much education – or thus human capital – he or she has acquired but on how much relative to others in the labour queue or, in other words, on his or her position in this queue (Thurow, 1983: 95-7).

If, in empirical analyses, education were to be treated on a relative basis, one outcome might then be that a weakening in the ED association was no longer so widely found. That is to say, an apparent weakening could be the result of sociologists implicitly thinking of education in absolute, ‘human capital’ terms while employers are in fact more often thinking in relative, ‘screening’ terms. From the few cases in which sociologists have so far sought to apply relative measures of education (Sørensen, 1979; Ultee, 1980; Wolbers et al., 2001) there are indications that a stronger ED association does in fact emerge than when an absolute measure is used, and further research on these lines might well be rewarding.

However, a relative perspective on educational attainment would seem likely to be still more illuminating in regard to the OE association. In this connection, a point that Thurow emphases (1975: 95-7) is that if it is an individual’s relative level of education that chiefly matters in the labour market, then pressure is likely to fall on parents and their children to ensure that the latter increase their educational attainment not just to acquire more human capital but, further, as an essentially ‘defensive’ measure: that is, in order to preserve their place within the labour queue – as, say, in the face of policies of educational expansion and reform. As well as this pressure being a possible source of further educational expansion having no necessary relation to demand and thus generating over-qualification (Chang, 2010: ch. 17), it could also result in more advantaged families taking the lead in processes of educational re-stratification, in which an expanding level of education becomes internally differentiated in terms of ‘quality’ (Lucas, 2001). And insofar as an educational ‘arms race’ is in this way set up, an explanation is suggested for the lack of clear evidence of a general and continuing weakening of the OE
association – and especially when educational re-stratification is adequately taken into account. The findings in this regard reported by Ichou and Vallet (2011) and by Boliver (2011) on persisting social inequalities in French secondary education and in British tertiary education, respectively, would seem entirely consistent with expectations under Thurow’s theory.

**Incentive-enhancing preference theory**

Incentive-enhancing preference theory is due to Bowles and Gintis (2000; see also Bowles, Gintis and Osborne, 2001a, b). Like screening and signalling theory and job competition theory, it is developed essentially in response to human capital theory. Bowles and Gintis do not deny that education brings an earnings return. However, they question whether education has this effect simply through increasing individuals’ knowledge and skills. In addition, they suggest, education plays a major role in socialising individuals for employment – that is, in shaping their values, norms and related preferences – and that in this way also it determines their potential productivity and thus their attractiveness as employees. Crucial in this regard is the recognition that Bowles and Gintis give to the distinctive nature of the employment contract: i.e. that it must to a significant extent be implicit and indeed incomplete. Because of this, it is important to employers that workers should have an orientation to work and ‘behavioural traits’ that will mitigate the problems that inevitably arise within the contract of worker effort and co-operativeness. For example, Bowles and Gintis pick out such traits as a high marginal utility of income, a low disutility of effort and a low rate of time preference (i.e. an orientation to the future rather than the present) which, they argue, are all likely to make workers more responsive to the incentives, and the sanctions, available to employers, and will therefore facilitate the ‘endogenous enforcement’ of the employment contract.
Bowles and Gintis acknowledge (2000: 125) that little research has been carried out into the role of schools in producing ‘incentive-enhancing preferences’, and that evidence in support of their theory is therefore mainly indirect (2001: 1147-56). They note that it has proved difficult to show that it is in fact knowledge and skills acquired in schools that actually generate earnings returns in the labour market; and they also point to findings from surveys of employers that suggest that the attributes of workers most in demand – and often seen as in short supply – are not ones relating to levels of formal qualification or even of skills but rather to ‘attitude’, ‘motivation’ or ‘personality’. However, empirical findings more directly relevant to their theory have in fact subsequently emerged. For example, work by Heckman and his associates (see e.g. Carneiro and Heckman, 2003; Heckman et al., 2006) indicates that a vector of non-cognitive traits of a kind likely to be incentive-enhancing does significantly raise earnings returns to schooling; and, more generally, findings are emerging (e.g. Osborne Groves, 2004; Jackson, 2006) on the importance of personality characteristics for occupational attainment, independently of formal qualifications.

Like screening and signalling theory and job competition theory, incentive-enhancing preference theory may appear somewhat over-generalised and in need of refinement in its application to different types of work. Bowles and Gintis see the problems of the employment contract, from the side of employers, as deriving from their principal-agent relationship with employees (e.g. 2000, p. 119). Contractual hazard arises in that the principal (the employer) engages an agent (the employee) to act in the principal’s interests in circumstances in which the principal cannot observe all of the agent’s actions nor have all of the information that is available to the agent, so that work monitoring is inherently difficult. This argument is clearly apposite in regard to professional employees engaged to exercise specialised knowledge and expertise, and to managerial employees to whom the employer delegates authority. However, it would seem rather strained to view in the same way the employer’s problems in regard to
employees who perform only low-skill, largely routine work. In this case, employers’ concerns are more likely to be with the quantity and predictability of the effort that is shown in performing work-tasks of a relatively well-defined kind (cf. Goldthorpe, 2007, vol. 2, ch. 5).

It could therefore be supposed that while some behavioural traits may be incentive-enhancing in a quite general way, the importance of others will be more specific to certain levels of employment. Bowles and Gintis do in fact at various points appear to recognise this. For example, they add to the list of ‘profitable’ traits previously mentioned those of leadership, initiative, ‘truth-telling’ and identification with organisational objectives, all of which would appear particularly relevant in the case of employees with whom a genuine principal-agent problem arises, while elsewhere they refer to such traits as dependability, consistency and punctuality, which would seem more relevant in the case of employees carrying out essentially routine activities.

However, the fact that incentive-enhancing preference theory does start out from a recognition of the problems inherent in the employment contract, which human capital theory ignores, gives it an obvious attraction from a sociological standpoint; and, with refinement on the lines indicated, the theory would appear to offer further potentially valuable insights into the social processes underlying empirical regularities observable within the OED triangle.

Thus, in regard to the ED association, what is of evident interest is how far, over and above imparting knowledge and skills, different types of educational institution, even at the same general level, may serve to socialise students into values, norms and preferences that will prepare them more or less well, from an employer’s point of view, for different grades of employment. A better understanding of the ED association, and of changes in it, might then be gained if mobility research were to take into account, in addition to the levels of
educational qualification that individuals attain, the amount of ‘seat time’ they have spent, over the course of their educational careers, in types of institution that could be supposed to differ in their ethos and in turn in the forms of socialisation in which they, knowingly or perhaps unknowingly, engage. Obvious institutional distinctions to make here would be private vs. state, religious vs. secular, single-sex vs. co-educational, selective vs. non-selective, and academic vs. vocational.

Research so far carried out focussing on such distinctions has been mainly concerned with effects on educational attainment per se. But this research does itself in some cases indicate that the advantages or disadvantages conferred on individuals in regard to their subsequent employment can extend beyond those associated with the formal qualifications they obtain. For example, in a study of later-life earnings of children educated in private and in state schools in Britain, Green et al. (2011) find that while a difference in earnings exists, in favour of the privately educated, and has been increasing, little more than half of it is attributable to their having superior qualifications.

Further, if socialisation for employment via education is considered in relation to socialisation within the family – a topic on which Bowles and Gintis have surprisingly little to say – there are relevant links to be made with well-known sociological research, in particular that of Kohn and his associates (see e.g. Kohn, 1969; Kohn et al., 1990) into the social-psychological processes through which distinctive orientations to work are intergenerationally sustained. This research shows that parents in professional and managerial positions affording them opportunities for self-direction and ‘ideational flexibility’ tend to emphasise these same attributes in the upbringing of their children, while parents in subordinate positions are more likely to emphasise qualities such as respect for authority, orderliness and reliability. In this way, then, a further possible source could be suggested of the OED interaction effect – if
interpreted as an effect of O on the ED association. In the case of children from more advantaged origins, the degree to which their family socialisation would tend to reinforce those aspects of their school socialisation favouring incentive-enhancing behavioural traits of the particular kind appropriate in higher-level personnel might be another way in which these children’s chances of advancement are increased even if their educational attainments are only modest. In contrast, for children of less advantaged origins family socialisation is likely to be more consistent with those aspects of school socialisation favouring behavioural traits more appropriate to routine employment, so that for these children the breaking of intergenerational continuity in level of employment becomes almost entirely dependent upon their educational attainment.

Conclusions

Social mobility research, and in particular research focusing on the role of education in mediating intergenerational mobility, for some time took as its theoretical context – insofar as its practitioners recognised any need for one – the functionalist theory of industrial and post-industrial society. However, the inadequacies of this theory have become increasingly apparent, and rather little appeal to it is now in fact made. Yet no alternative body of theory has been developed to inform research, with the result that findings going contrary to the temporal and cross-national regularities that the functionalist theory would predict are themselves left largely unaccounted for. I have proposed that in developing the micro-social foundations that are needed for a better understanding of macro-level relations within the OED triangle, sociologists could, with advantage, draw on theories originating in the economics of labour markets and, in particular, on three theories in some degree deviating
from mainstream human capital theory: i.e. screening and signalling theory, job competition theory and incentive-enhancing preference theory.

I have aimed to substantiate this proposal by showing how, in the light of these theories, explanations can be put forward for the widespread tendency for the ED association to weaken rather than to strengthen - at least where education is understood in terms of absolute level of education or qualification achieved; for the absence of any general and sustained weakening in either the OE or the OD associations; and for the interaction effect typically found within the OED triangle, at least on one interpretation of this effect. I have also in most cases been able to show that the possible explanations advanced are not entirely a matter of speculation but are consistent with various extant findings; and, more importantly, I have been able to point to ways in which research aimed at the more specific empirical testing of these explanations might be pursued.

As regards research into the role of education in social mobility more generally, two points of importance emerge. First, in mediating the mobility chances of individuals, education should not be seen as operating in some single mode. Rather, it has to be recognised that it may operate in several, quite different modes: as a source of human capital and thus of actual productive capacity; as a basis - along with other individual attributes – for the screening and signalling of productive potential; as a positional good determining individuals’ chances in processes of job competition; and as an agency of socialisation endowing individuals with values, norms and preferences more or less appropriate, from an employer’s point of view, for different levels of work. Second, it seems likely that these different modes will be of differing relative importance – as also may be the role of education itself, however viewed – according to economic sector and to type and level of employment.
Two corresponding methodological requirements can then be identified. First, the variable of ‘education’ should be conceptualised and measured in differing ways. Due attention must be given to whether educational qualifications are being viewed in a certifying or in a screening and signalling role, and to whether education should be treated in terms of relative as well as of absolute attainment, so that individuals can be positioned in the overall distribution of education within a population or birth cohort. In addition, it may also be important to treat education according to differences in the institutions in which it has been obtained insofar as these may be relevant to processes of socialisation for working life as well as to the acquisition of knowledge, skills and formal qualifications. Second, standard analyses of the role of education in mediating social mobility at the aggregate population level need to be complemented by more disaggregated analyses, by sector and type and level of employment, so that a better understanding can be gained of the quantitative significance of its different possible modes of operation and of their likely compositional effects.
References


Klein, M. (2011) ‘Trends in the association between educational attainment and class destinations in West Germany: Looking inside the service class’. Research in Social Stratification and Mobility, 29, 427-44.


