

Breadth versus Depth: The Timing of Specialization in Higher Education

Ofer Malamud

Abstract. This paper examines the trade-off between early and late specialization in the context of higher education. I develop a model in which individuals accumulate field-specific skills and receive noisy signals of match quality across different fields of study. I derive comparative static predictions between educational regimes with early and late specialization, and examine these predictions across British systems of higher education. Using survey data on 1980 university graduates, I find that individuals who switch to unrelated occupations have lower initial earnings, and that early specialization in England is associated with more costly switches. But higher wage growth among those who switch eliminates the wage difference after several years, and average earnings are not significantly different between England and Scotland.

1. Introduction

Division of labor — the tendency of individuals to specialize in specific occupations — is an important feature of the modern labor market. However, for many professional occupations, such as those held by scientists, engineers, managers, lawyers, and teachers, specialization begins prior to labor market entry when an individual chooses a field of study in university. The timing of such academic specialization varies across different systems of higher education. In some systems, students are required to choose a field of study before they apply to college. In others, students may postpone the decision until late in their college careers. These differences highlight the trade-off between accumulating more human capital in a particular field by specializing early versus gathering additional information about alternative fields by specializing later. I explore the consequences of early and late specialization by comparing wages and other labor market outcomes across two educational systems with different exogenous constraints on the timing of academic specialization.¹

Ofer Malamud, Harris School of Public Policy, University of Chicago, Chicago, IL 60637, USA.
E-mail: malamud@uchicago.edu.

I wish to thank Claudia Goldin, Caroline Hoxby, Larry Katz, Franco Peracchi, and an anonymous referee for extensive comments, as well as Itai Ashlagi, Nittai Bergman, Saar Golde, Jeff Grogger, Michael Kremer, Seema Jayachandran, Bob Lalonde, Steve Machin, Klaus Miescke, Derek Neal, Steve Pischke, Cristian Pop-Eleches, Sarah Reber, Bruce Sacerdote, Anna Vignoles, Abigail Waggoner, Tara Watson, and seminar participants at Clemson University, Hebrew University, Harvard University, Michigan State University, Tel-Aviv University, UCSD, UC River-side, University of Chicago, and the NBER Higher Education meeting for many helpful suggestions. I am grateful to the Universities Statistical Record, the UK Data Archive, and several university administrators in Scotland and England for assistance. All errors are my own. This work was supported by a grant from the Spencer Foundation.

I introduce a simple model in which individuals initially take courses in a number of different fields of study but specialize at some point by choosing one field and taking their remaining courses in this field exclusively. A key aspect of the model is that individuals both accumulate field-specific skills and learn about their unobserved match quality to different fields over time. Later specialization provides students with more time to learn about match quality in different fields but it affords less time to acquire field-specific skills after a field is chosen. Assuming that wages are increasing in both field-specific skills and match quality, I show that later specialization is optimal when the return to match quality is high relative to the return to specific skills. Extending the model to allow for switching to occupations, which are unrelated to one's chosen field of study, I predict that individuals who switch fields will earn lower wages than those who enter related fields. This is because switching is associated with a loss in specific skills and because match quality conditional on switching is, on average, lower. Moreover, since switching allows individuals to correct for poor choices made at the point of specialization, it should be more valuable in regimes where individuals are required to specialize early. Consequently, the difference in expected wages between the early and late regime should diminish when switching is possible and the return to match quality is relatively high.

I proceed to examine the labor market consequences of specializing early versus late by comparing across the English and Scottish undergraduate systems. In England, students apply to a specific field of study at a particular university while still in secondary school. Once admitted to study a certain field, they usually follow a narrow curriculum that focuses on the chosen subject and allows for few courses in other fields. That is, English students are required to specialize early. In contrast, Scottish students are typically admitted to a broad faculty rather than a specific field. They are required to take several different subjects during their first two years before specializing in a particular field. That is, Scottish students are required to generalize early and specialize late. With a few exceptions, these differences in the timing of academic specialization between England and Scotland have existed for more than a century.² I focus on the comparison between England and Scotland because, although their educational systems are separate and arguably exogenously different, their labor markets are relatively well integrated and macroeconomic policies are determined by a common government.

Using survey data on 1980 university graduates from England and Scotland, I find strong evidence that individuals who switch to unrelated occupations earn lower initial wages. This confirms one of the main predictions of the model.³ Although imprecise, estimates of wage differentials between switchers and non-switchers indicate that the cost of switching may be higher in England than in Scotland. Nevertheless, individuals who switch also experience greater wage growth so that most of the wage differential becomes insignificant after 6 years in the labor market. Furthermore, as the likelihood of switching fields is substantially higher in England than in Scotland, the model suggests that students in England make more mistakes in choosing their field of study and suffer a corresponding loss in specific skills when trying to correct them. Finally, controlling for demographic and occupational characteristics, there is no significant difference in average wages and reported subjective satisfaction between individuals in England and Scotland. This is consistent with the prediction that differences in expected wages between the England and Scotland narrow when the return to match quality is relatively high. Overall, these findings suggest that early specialization is associated with more costly switches upon entering the labor market but that these differences between early and late specialization do not persist over time.

The model of academic specialization presented in this paper is related to the distinction between general and specific education. For example, Wiswall (2008) introduces a model where teacher licensing requires students to specialize in teaching earlier, reducing the

accumulation of general skills, which are useful for those who decide not to teach. Similarly, in my model, an educational system with early specialization provides for more specific skills in a particular field but fewer skills in a range of fields, as compared with an educational system with later specialization. Thus, later specialization would tend to yield better outcomes in the presence of substantial labor market volatility.⁴ But my model also allows for imperfect information about match quality. In a static labor market with perfect information, Weiss (1971) shows that it is not optimal to delay the investment in education or change occupations. Allowing for imperfect information about match quality, I find that it may be better to delay specialization even in a relatively static labor market. Furthermore, the arrival of new information about match quality can lead some individuals to switch to an occupation that is unrelated to their chosen field of study.

That education may provide individuals with valuable information about their match quality to different fields of study is not emphasized much in the literature on human capital.⁵ Some notable exceptions include Altonji (1993), who introduces a model where individuals learn their preference between two fields of study by attending college, and Arcidiacono (2004) who estimates a structural model of student learning. The theoretical framework presented in this paper embeds both skill acquisition and learning within a model of academic specialization. Malamud (2011) uses a similar framework to show that higher education, in addition to providing skills, also provides information about one's tastes and talents for different fields of study. The present paper derives and tests a new set of reduced-form predictions regarding wages both within and between regimes with early and late specialization. Note that the process of learning about match quality in a particular field is complementary to the acquisition of specific skills in that field. Thus, in contrast to the competing tasks of on-the-job search and firm-specific human capital acquisition in Jovanovic (1979b), the trade-off associated with academic specialization arises *not* between the accumulation of human capital and learning about match quality *per se*, but rather, between the accumulation of human capital in a *particular field* and the possibility of learning about match quality in *alternative fields*.

The paper proceeds as follows. Section 2 develops a simple model of academic specialization and derives comparative static predictions across regimes with early and late academic specialization. Section 3 extends the model to allow for switching to occupational fields, which are unrelated to fields of study. Section 4 explores the differences between the English and Scottish systems of higher education in more detail. Section 5 describes the data and the empirical methodology. Section 6 presents results from the regression analysis. Section 7 concludes.

2. A simple model of academic specialization

2.1 Set-up

Suppose individuals take n courses in each of k fields of study prior to specialization. At the point of specialization, individuals choose a single field and take $(N - nk)$ additional courses in this chosen field of study. After completing a total of N courses, individuals enter an occupation in their chosen field. Assume that individuals are risk neutral and have identical prior distributions on match quality for each field. Prior to undertaking schooling, however, match quality in each field is unknown. Specifically, assume that match quality, θ_i , in each field i is a random draw from a normal distribution with the same mean and variance, so that $\theta_i \sim N(\mu, \sigma_0^2)$. Match quality is therefore uncorrelated across fields. Match quality can include any field-specific component of education that affects wages, such as inherent ability or

interest in the subject matter. Allowing for prior means and variances to differ across fields is straightforward and does not alter the main results from the model so long as we abstract from the possibility of switching fields later on.

By taking courses in a given field, individuals will (i) accumulate field-specific skills and (ii) receive noisy signals of their match quality in that field. For simplicity, suppose that the quantity of skills accumulated in a field, s_i , is equivalent to the number of courses spent studying that field. Each course of study j in field i provides a signal of match quality in that field, $x_{ij} = \theta_i + \varepsilon_{ij}$ where $\varepsilon_{ij} \sim N(0, \sigma^2)$ and $j = 1, \dots, n$. Noise in the signal may be due to any number of idiosyncratic factors such as the quality of instruction or the particular circumstances of the student at the time. I assume that skills are perfectly specific to a particular field but I discuss the possibility of spillovers across fields in Section 2.3.

Upon entering the labor market, match quality is revealed. The wage in field i is an increasing function of both match quality and skills: $w_i = w(\theta_i, s_i)$ so that $\partial w/\partial \theta > 0$ and $\partial w/\partial s > 0$. For simplicity, I assume that wages are a linear function of match quality and skills, $w(\theta_i, s_i) = \alpha\theta_i + \beta s_i$. I take (α/β) as an indication of the return to match quality relative to the return to specific skills. More generally, we might expect a different functional form for wages across different fields.⁶ In the empirical analysis, I compare outcomes for individuals controlling for field of study to account for mean differences in wages across fields. Finally, for the purposes of the empirical analysis, I suppose that individuals only consider wages when making educational and occupational decisions. However, if instead, I were to consider utility as a function of wages as well as non-pecuniary factors, I would derive analogous predictions for utility.

2.2 Choice of field at specialization

The posterior distribution of match quality after studying n courses in field i is a normal distribution with mean μ'_i and variance σ' .⁷ And the quantity of skills in each field at the point of specialization is $s' = n$. Therefore, in specializing, risk-neutral individuals with identical prior distributions across fields will choose the field of study with the highest expected wages:

$$\text{choose } i^* = \arg \max_{i=1, \dots, k} \left\{ E \left[w(\theta_i, s_i) \mid \{x_{ij}\}^{j=1, \dots, n} \right] \right\} = \arg \max_{i=1, \dots, k} \{ \alpha \mu'_i + \beta s' \}.$$

As the quantity of specific skills in each field is identical, individuals simply choose the field with the highest posterior mean of match quality, $i^* = \arg \max_{i=1, \dots, k} \{ \mu'_i \}$.⁸ Thus, the posterior mean of match quality in the chosen field at the time of specialization will be μ'_{i^*} .⁹ Introducing risk aversion does not alter the decision at the point of specialization if the variances of the prior distributions across fields are identical; individuals would continue to choose the field with the highest posterior mean. However, if more precise information is available about certain fields at the point of specialization (i.e. σ_0^2 varies by field), risk-averse individuals could decide to choose such fields even when they are associated with lower posterior means.

2.3 Optimal timing of specialization

Suppose that no field switching is permitted so that individuals must enter their chosen field of study. Individuals who specialize later have less time to accumulate specific skills in their chosen field of study but receive more signals in each field prior to specialization. They will therefore have more accurate assessments of their match quality in each field and be less likely to make a mistake in choosing a field. Thus, the optimal point of specialization depends on the return to match quality relative to the return to specific skills.

Proposition 1: The optimal number of courses prior to specialization, n , is increasing in $\alpha\beta$.

See Appendix B for a formal proof. Now consider predictions on wages for regimes with early and late specialization: An early regime requires individuals to specialize after taking n^E courses in each field; a late regime requires individuals to specialize after taking n^L courses in each field, where $n^E < n^L$. As before, specific skills will be lower and match quality will, on average, be higher for individuals in the late regime. Hence, whether individuals in the early regime ultimately earn higher expected wages than their counterparts in the late regime will depend on the return to match quality relative to the return to field-specific skills.

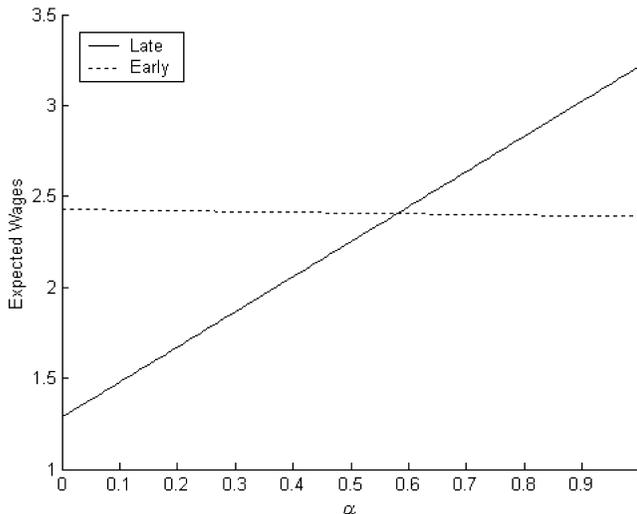
Corollary 1: A regime with late specialization, n^L , will have higher wages than a regime with early specialization, n^E , if the return to match quality is sufficiently higher than the return to specific skills:

$$E[w^L(\theta_{i^*}, s)] > E[w^E(\theta_{i^*}, s)] \Leftrightarrow \frac{\alpha}{\beta} > \Omega > 0.$$

Appendix B provides a proof. Simulations of expected wages also show the behavior of wages over a broad set of parameter values.¹⁰ Figure 1 plots expected wages for an early and a late regime over the full range of relative returns to match quality, which are normalized by taking $\beta = (1 - \alpha)$ so that (α/β) goes from 0 to ∞ as α goes from 0 to 1. When the relative return to match quality is high, individuals who specialize later will earn higher wages.

As mentioned earlier, skills are assumed to be specific to a particular field. Allowing for spillovers across fields would serve to dampen the trade-off between match quality and skills as additional learning about match quality would be less costly in terms of forgone skill acquisition. Moreover, it is possible to incorporate a notion of general skills more directly: for example, by letting wages also depend on average skill in the fields not chosen for specialization, $g = (1/k)\sum_{i \neq i^*} s_i$, or on the minimum level of skills in any field $g = \min_{i=1, \dots, k} \{s_i\}$. This

Figure 1. Expected wages without field switching by relative return to match quality



would serve to increase the benefits associated with later specialization but the qualitative predictions of Proposition 1 and Corollary 1 would remain unchanged. More generally, it is possible to characterize educational regimes according to whether they provide more general skills or more specific skills. In this case, the question of which system is better would depend on the relative returns to general versus specific skills. Although this can generate similar predictions regarding average wages between the two regimes, it does not generate differences in the likelihood of switching.¹¹

3. Academic specialization with field switching

3.1 Decision on whether to switch

Now suppose that individuals can switch to an occupational field, which is unrelated to their field of study prior to entering the labor market. Following specialization, individuals take $(N - nk)$ additional courses in the chosen field and receive more signals about match quality in the chosen field, i^* . The posterior distribution of match quality in the chosen field after $(N - nk)$ additional signals will be updated to a normal distribution with mean μ_{i^*}'' and variance σ'' . Moreover, the quantity of skills in the chosen field prior to entering the labor market is $s'' = n + (N - nk)$. So now, given the opportunity to switch to another field prior to entering the labor market, individuals will compare expected wages in the chosen field with expected wages in the next best field:

$$\begin{aligned} \text{field switch} &\Leftrightarrow E[w(\theta_i, s_i) | \{x_{ij}\}^{j=1, \dots, N}] < \max_{i \neq i^*} \{E[w(\theta_i, s_i) | \{x_{ij}\}^{j=1, \dots, n}]\} \\ &\Leftrightarrow \alpha \mu_{i^*}'' + \beta s'' < \max_{i \neq i^*} \{\alpha \mu_i' + \beta s'\}. \end{aligned}$$

Intuitively, individuals will switch if the posterior mean of match quality in the chosen field falls sufficiently far below the posterior mean of another field to overwhelm the loss in specific skills from switching. If individuals decide to switch, they will always choose the field with the second highest posterior mean since all fields other than the one chosen are associated with the same quantity of specific skills and posterior variance. The decision of whether to switch can therefore be framed as a comparison between the first best field, i^* , and the field that was second best at the time of specialization, i'' . The field selected after the second stage is denoted i^{**} where $i^{**} \in \{i^*, i''\}$.

The probability of switching to an unrelated occupational field depends on the timing of specialization. Whether the probability of switching is higher in a regime with early or late academic specialization also depends, in turn, on the return to match quality relative to the return on specific skills. In an early regime, assessments of perceived match quality in the chosen field experience relatively greater updating following specialization so individuals are more likely to realize they made a mistake and hope to correct it by switching. However, individuals in an early regime also lose more specific skills by switching fields. Hence, the probability of switching will be higher in an early regime when the relative return to match quality is sufficiently high.¹²

3.2 Choice of field at specialization

If the variance of priors on match quality are identical across fields, individuals do not need to consider the possibility of later switching when making their initial choice of field. However,

allowing prior variances on match quality to vary by field introduces option value considerations at the time of specialization. Similar to the prediction derived by Miller (1984), individuals would then tend to specialize in riskier fields because they could switch in case of a bad realization.¹³ Moreover, fields with a larger prior variance would have greater option value in the early regime than in a late regime. With more signals following specialization, greater updating in an early regime generates a higher probability that the ultimate posterior mean will surpass that of the chosen field. Hence, individuals in an early regime will be more likely to choose a field with a lower posterior mean at the point of specialization because of the greater option value. As, on average, such fields have lower expected match quality than those with the highest posterior mean, we expect more field switching in an early regime due to option value considerations. Finally, note that fields of study are assumed to provide only field-specific skills. Therefore, the model does not necessarily predict that individuals will choose a different set of fields in early and late regimes.

3.3 Wages

Allowing for switching, it is possible to characterize expected wage differences between those who switch and those who do not switch *within* regimes. The quantity of specific skills for individuals who switch to occupations unrelated to their chosen field of study is always lower than for those who enter related occupations. Furthermore, match quality conditional on switching is lower in expectation since it is chosen fields with lower match quality that ultimately lead to bad signals and cause switching. Thus, on average, individuals who switch will have lower levels of both match quality and specific skills than those who do not switch.¹⁴

Proposition 2: Individuals who switch will have lower wages than those who do not switch:

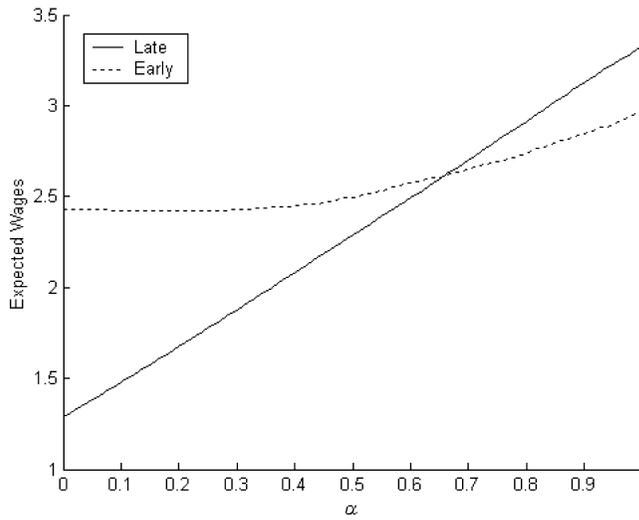
$$E[w(\theta_{i^*}, s) | w(\theta_{i^*}, s) > w(\theta_{i^a}, s)] - E[w(\theta_{i^a}, s) | w(\theta_{i^a}, s) > w(\theta_{i^*}, s)] > 0.$$

See Appendix B for a formal proof. As an extension, suppose that individuals continue to accumulate field-specific skills on the job, either from on-the-job training or through learning by doing. Then, if there are diminishing returns to specific skills, individuals who switch will have higher rates of wage growth as they begin with lower levels of specific skills in their occupational fields.

The possibility of switching also affects expected wage comparisons across regimes because individuals have the opportunity to correct their initial choices. Indeed, on average, those individuals who switch earn higher wages than they would have earned in the baseline case without switching. As more mistakes are made with early specialization, it is more valuable to be able to correct them through switching in the early regime. Consequently, we expect the difference in expected wages between the early and late regime to be dampened when the return to match quality is relatively high. Nevertheless, the underlying intuition for expected wage differences is analogous to the Corollary of Proposition 1; when the return to match quality is relatively high, later specialization is optimal because it provides for more signals in each field prior to specialization. Thus, we have the following proposition.

Proposition 3: With field switching, a regime with late specialization, n^L , will have higher wages than a regime with early specialization, n^E , if the return to match quality is sufficiently higher than the return to specific skills:

Figure 2. Expected wages with field switching by relative return to match quality



Notes: All simulations are based on 5,000 repetitions for $k = 2$, $N = 21$, $\mu = 0$, $\sigma_0 = 25$, and $\sigma = 100$. Early regimes are characterized by $n^E = 2$; late regimes are characterized by $n^L = 6$. The relative returns to match quality are normalized by taking $\beta = (1 - \alpha)$ so that (α/β) goes from 0 to ∞ as α goes from 0 to 1. Expected wages are log wages determined according to $E(\ln w_j) = E(\alpha\theta_i + \beta s_i)$ where $s_j = [s_j/(N/k)] + \mu$ are normalized skills.

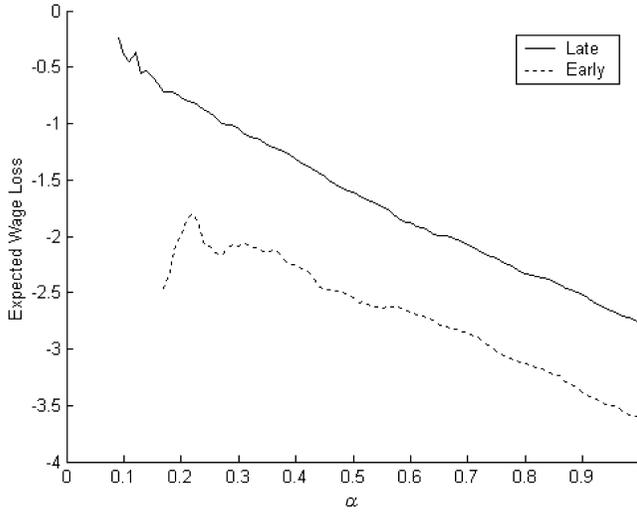
$$E[w^L(\theta_{i^{**}}, s)] > E[w^E(\theta_{i^{**}}, s)] \Leftrightarrow \frac{\alpha}{\beta} > \Omega' > 0.$$

See Appendix B for further discussion. Expected wages in an early and a late regime are shown in Figure 2 for different relative returns to match quality. Note that expected wages in the early regime converge toward expected wages in the late regime as the relative return to match quality rises and more switching takes place. As in the case without switching, this prediction indicates that the superiority of one regime over the other depends critically on the relative returns to match quality and field-specific skills.

Finally, consider the difference in the average wage loss due to switching between an early and late regime. The difference in specific skills between individuals who do not switch and those who switch will always be greater in the early regime than in the late regime as $(s'' - s')$ is decreasing in n . However, individuals in the late regime will generally have higher match quality in both the chosen and the alternative fields of study as they receive more signals on match quality in all fields prior to specialization. Consequently, the overall difference in the expected wage loss from switching is ambiguous. In the case of switching for exogenous reasons unrelated to match quality, the intuition is clear. Individuals in the early regime make more mistakes in their initial choice of field at the time of specialization so random switches are more likely to yield corrections to higher match quality in the early regime than in the late regime. Thus, with a sufficiently high return of match quality relative to specific skills, the average wage loss will be greater in the late regime.

Proposition 4: With exogenous switching, a regime with late specialization, n^L , yields a larger average expected wage loss from switching than a regime with early specialization, n^E , if the return to match quality is sufficiently higher than the return to specific skills.

Figure 3. Expected wage loss by relative return to match quality



Notes: All simulations are based on 5,000 repetitions for $k = 2$, $N = 21$, $\mu = 0$, $\sigma_0 = 25$, and $\sigma = 100$. Early regimes are characterized by $n^E = 2$; late regimes are characterized by $n^L = 6$. The relative returns to match quality are normalized by taking $\beta = (1 - \alpha)$ so that (α/β) goes from 0 to ∞ as α goes from 0 to 1. Expected wages are log wages determined according to $E(\ln w_j) = E(\alpha\theta_i + \beta s_i)$ where $s_i = [s_i/(N/k)] + \mu$ are normalized skills.

See Appendix B for a formal proof. For the case of endogenous switching described in this model, those individuals who switch are precisely the ones that receive bad signals after specialization. As a result, individuals who would otherwise have tended to earn low wages as non-switchers will tend to earn higher wages as switchers and the expected wage loss from switching will be attenuated. This attenuation will generally be greater in the late regime because more erroneous switches are avoided and implies that the expected wages loss may actually be larger in the early regime even when the return to match quality is high.¹⁵ Figure 3 shows the wage differential in an early and a late regime for a range of relative returns to match quality for the case of endogenous switching.

4. Background: higher education in Britain

The British system of higher education provides a particularly appropriate setting in which to examine the predictions of the model. Undergraduate education in England and Scotland, although similar in aim and overall structure, varies widely in the timing of academic specialization. In England, students apply to a specific field of study at a particular university.¹⁶ Once admitted to a specific field, English students usually follow a narrow curriculum that focuses on the main field and allows for little exposure to other fields.¹⁷ Indeed, most universities in England require students who change fields of study to start university anew (although some do allow for limited changes). In contrast, Scottish students are typically admitted to a broad faculty or school rather than a department; in some universities, admission is to the university at large.¹⁸ Furthermore, they are required to study several different fields during their first two years. As an undergraduate prospectus for the University of Edinburgh (2003) explains:

You would normally take courses in three or more subjects in the first year and, commonly, these are followed by second courses in at least two of the subjects in your second year. This will then

give you a choice from two, or even three, subjects to pursue to degree level, and you can delay this decision until quite a late stage . . . In choosing courses to be taken in the first two years, you can select from a very wide range of courses offered across several faculties.

Similar course structures exist in most Scottish universities. Scottish universities thus allow for substantial choice among fields of study within faculties and, to some degree, across faculties as well.¹⁹ Moreover, students in Scotland are *required* to take a broader range of courses and choose a field of study much later than their English counterparts.²⁰ The Handbook for Students and their Advisors of 1980–82 explains that ‘the standard English degree, whether in science, humanities or social sciences, is a single subject honours degree’ whereas ‘universities in Scotland had traditionally offered a wide range of subject options with multi-subject examinations at the end of the first year’ (pp. 17–18). This is also supported by empirical evidence that the proportion of individuals who change their field of study between admission and graduation in Scottish universities is substantially higher than in English universities (see Malamud, 2011). Given these differences, it is quite natural to regard the English system of higher education as an early regime and the Scottish system of higher education as a late regime.

There is variation in the average length of the undergraduate degree between England and Scotland. Although there is some heterogeneity among degrees within each nation, most English degrees are completed within 3 years whereas most Scottish degrees are completed within 4 years. However, many Scottish students enter university after 6 years of secondary schooling rather than the 7 years customary in England. According to this calculation, English and Scottish students who attain a BA degree receive roughly the same number of years of schooling (and this is confirmed in the data by examining the age of graduation). Loosely speaking, the first year of university in Scotland may be said to correspond to the final year of secondary school in England. But even so, as English students apply to university in the beginning of their final year of secondary school whereas Scottish students only make their final choice of field at the end of their second year of university, there is substantial difference in the timing of specialization.

The difference between English and Scottish universities arose from their unique respective historical traditions. English universities were largely independent and free to set their curriculum and course structures. The provincial civic universities established later in urban centers did not substantially depart from the traditions of the ‘ancient’ universities. Even with the introduction of broad faculties and additional courses of study, admissions remained at the departmental level.²¹ On the other hand, Scottish universities became regulated under the Universities (Scotland) Act of 1858 that set up an executive commission to draw up uniform conditions for courses of study. The Universities (Scotland) Act of 1889 further increased the choice of subjects available in Scottish universities, reflecting the ‘traditional Scottish preference for a broad general education’ (Hunter, 1971, p. 237). In large part, these two Acts of Scottish Parliament determined the distinctive characteristics of universities in Scotland, including the emphasis on late academic specialization.

In addition to differences in higher education, England and Scotland also differ in their system of secondary school education. In England, students need General Certificate of Education (GCE) Advanced-level examinations (A-levels) in two or three subjects to gain acceptance into university.²² In 1989, a new exam, the Advanced Supplementary examination (AS-level), was brought in to broaden the curriculum; it was to be the same standard as an A-level, but half the content. Students were encouraged to substitute two AS-levels for one of their A-levels but most universities did not regard these examinations as commensurate alternatives and it did little to change the character of English secondary school education. In

Scotland, on the other hand, students need Scottish Certificate of Education (SCE) Higher Examinations in five or six subjects to gain acceptance into university.²³ More recently, Advanced Highers and Higher Still certifications have been introduced to provide the opportunity for further specialization in secondary school. However, universities continue to use Highers as the primary basis for admission and there is little doubt that the Scottish system of secondary education provides a broader curriculum than the English one. Again, the reasons for these differences in secondary school curriculum can be traced to historical antecedents. In effect, specialization trickled down from the universities to secondary schools. Moreover, the early influence of English universities on secondary school leaving exams was far stronger than that of Scottish universities since Scottish secondary school leaving certificates had to be approved by the Scottish Education Department.

5. Data and empirical strategy

5.1 Data

Data for the empirical analysis come from the 1980 National Survey of Graduates and Diplomates (NSGD). The NSGD was a national postal survey of some 4,800 university graduates undertaken in 1986/87 by the British Department of Employment.²⁴ The NSGD contains information about their 1980 qualification, their subsequent labor market experience (occupation, industry, and earnings for first and current job), and further educational pursuits. There is also information about their high school examination results and some questions regarding satisfaction with the 1980 qualification. Although it is not possible to identify specific universities in the NSGD, there is information on whether students took English or Scottish secondary school leaving exams. Indeed, using school leaving exams as a proxy for type of degree serves to reduce the bias associated with non-random migration to university.²⁵ As the NSGD is not representative of the overall population, we might be concerned that the English and Scottish samples of university graduates may not be comparable because of differing enrollment rates. Hence, I use two nationally representative data sets, which include all individuals born in Great Britain during 1 week in 1958 and 1970 (the National Child Development Study and British Cohort Study, respectively), to calculate the percentage of individuals who have attained a first degree from university by age 26. In both data sets, enrollment rates in university are remarkably similar between England and Scotland: 8 per cent of the 1958 cohort and 12 per cent of the 1970 cohort.

Table 1 shows the average characteristics for the sample of English and Scottish students used in the regression analysis.²⁶ The average age upon completion of the first degree is almost equivalent among English and Scottish students. Although the average age that students begin university is slightly lower in Scotland, the median age of students during their first year in university is 19 for both England and Scotland (not shown). The raw high school grade point average (GPA) shown in Table 1 is converted from letter grades in the A-level and Scottish Higher school leaving examinations. In the regression analysis, these scores are normalized within nation so that coefficients represent the effect of a one standard deviation increase in GPA. The composition of broad fields of study across the two nations is not too dissimilar, especially after accounting for the oversampling of engineering students from Scotland. Nevertheless, relatively more students in Scotland study life sciences, health sciences, and business and relatively fewer study mathematical and social sciences. The majority of students from England and Scotland enter employment in the UK. The lower rate of unemployment among

Table 1. Summary statistics for 1980 college graduates

	England			Scotland		
	Mean	SD	Observations	Mean	SD	Observations
<i>Individual characteristics</i>						
Female	0.34	0.47	1,242	0.31	0.47	213
Married (6 years after degree)	0.53	0.50	1,242	0.59	0.49	213
Average age (upon completion)	22.01	1.51	1,242	22.26	2.40	213
High school GPA (out of 30)	19.71	5.84	1,242	18.25	5.77	213
Number of high school subjects	3.18	0.69	1,242	5.15	1.04	213
<i>Degree characteristics</i>						
Math and Computer Sciences	0.08	0.26	1,242	0.04	0.19	213
Physical Sciences	0.15	0.35	1,242	0.08	0.26	213
Architecture	0.02	0.13	1,242	0.02	0.15	213
Engineering	0.21	0.41	1,242	0.30	0.46	213
Life Sciences	0.07	0.25	1,242	0.08	0.27	213
Health Sciences	0.04	0.20	1,242	0.05	0.22	213
Social Services and Welfare	0.03	0.17	1,242	0.02	0.14	213
Social Sciences	0.19	0.39	1,242	0.15	0.36	213
Business/Accounting	0.04	0.21	1,242	0.06	0.24	213
Law	0.01	0.11	1,242	0.10	0.30	213
Education	0.03	0.17	1,242	0.04	0.20	213
Art	0.13	0.34	1,242	0.06	0.23	213
<i>Field switching</i>						
Very broad classification	0.44	0.50	1,242	0.29	0.45	213
Broad classification	0.50	0.50	1,242	0.34	0.48	213
Narrow classification	0.63	0.48	1,242	0.51	0.50	213
<i>Labor market outcomes</i>						
Log earnings after 1 year	8.43	0.32	1,173	8.38	0.37	206
Log earnings after 6 years	9.12	0.47	1,173	9.09	0.42	206
<i>Region of work</i>						
England	0.87	0.33	1,242	0.25	0.44	213
Scotland	0.02	0.13	1,242	0.71	0.46	213
Wales	0.03	0.18	1,242	0.00	0.07	213
Northern Ireland	0.00	0.06	1,242	0.00	0.00	213
Abroad	0.07	0.26	1,242	0.03	0.18	213
<i>Additional outcomes (full sample)</i>						
Entering employment ^a	0.62	0.49	3,663	0.64	0.48	538
Further study ^a	0.28	0.45	3,663	0.29	0.45	538
Unemployed ^a	0.10	0.31	3,663	0.07	0.25	538

Notes: The base sample for the 1980 NSGD includes all individuals who attained a BA degree in 1980 and were employed in a job during the first year following graduation and not pursuing graduate studies. Median age at the start of the degree is 19 for both nations. GPA is an average measure of the achievement in secondary school leaving exams out of 30 (but standardized by nation in all regressions). Field switching is defined as 1 if field of study at the undergraduate level is different from the occupational field of first job 6 months following degree and 0 otherwise (see Appendix A for further discussion of classification groups). Log earnings are expressed in 1981 pounds after deflating using the consumer price index. Composition of fields of study and occupational fields are based on a broad classification (other classifications are discussed in Appendix A). Foreign students returning overseas are excluded from counts of Post-BA activity.

^a is out of the unrestricted sample including unemployed and graduate students.

Scottish individuals is a consequence of the oversampling of engineering graduates who are less likely to be unemployed than others.²⁷

The model introduces an important distinction between individuals who enter an occupation that is related to their field of study and those who switch to an unrelated occupation. I

construct a variable *SWITCH* that captures field switching by grouping fields of study and occupations into categories (see Appendix A for more details). As shown in Table A1, I allow for three levels of classification: narrow (42 categories), broad (12 categories), and very broad (6 categories). Individuals are said to switch to an unrelated occupation when the field of study of their degree and their occupational field are in different categories, subject to the level of classification. Therefore, a field switch is defined as 1 if the occupational field is different from the field of study at university, and 0 otherwise. Broader classifications indicate lower rates of field switching as only drastic changes from fields of study to occupational fields will register. However, the rate of field switching is substantially lower in Scotland than in England according to all classifications. For example, in terms of the broad classification, the rate of field switching in Scotland is between 10 and 20 percentage points lower than the rate of field switching in England. Most of the empirical analysis will focus on the broad classification of fields.²⁸

5.2 Empirical strategy

The base sample includes all individuals who attained a BA degree in 1980 and were employed full-time in the first year following completion of their qualification. I exclude individuals pursuing graduate studies while working because this may select for weaker students who need to work while pursuing higher degrees. I explore a variety of alternative sampling restrictions: (i) including graduate students who have occupation data; (ii) including unclassified occupations such as manual and clerical occupations instead of coding them as switches since individuals in one nation may be more likely to end up in non-professional occupations; (iii) coding individuals who end up unemployed as switches since this may be the result of a differential macroeconomic shock across the two nations; and (iv) excluding the fields of education and business or coding individuals who study them as non-switches since they are particularly subject to misclassification (and similarly with combined fields). Finally, I check that the findings hold for students with top high school grades who are clearly free to choose their fields, unconstrained by admissions requirements and the availability of slots.

The theoretical predictions regarding wages are examined through the following regression model:

$$\ln w_{ij} = \beta' \mathbf{X}_{ij} + \lambda SCOT_{ij} + \gamma SWITCH_{ij} + \delta(SCOT_{ij} \times SWITCH_{ij}) + \phi_j + \varepsilon_{ij}, \quad [1]$$

where $\ln w_{ij}$ is log annual wages or earnings for individual i in field j , $SCOT_{ij}$ is a dummy variable indicating the individual received a Scottish degree and therefore specialized late, $SWITCH_{ij}$ is a dummy variable for a field switch, ϕ_j is a set of field of study effects, \mathbf{X}_{ij} are demographic characteristics, and ε_{ij} is a disturbance term. The main demographic controls include sex, age, marital status, high school GPA, parental socioeconomic status, as well as controls for region of work and industry. Since the log function is a positive monotonic transformation, all of the predictions derived in Section 3 on wages will also hold for log wages. λ captures the difference in wages between England and Scotland among individuals who do not experience field switching. γ captures the differential in wages in England between individuals who switch and those who do not switch. Finally, δ captures the *difference* between Scotland and England in the differential associated with switching. Other parameters of interest include the wage differential from switching for individuals in Scotland ($\gamma + \delta$) and the wage difference between English and Scottish individuals who switch ($\lambda + \delta$). All wage regressions use the type of high school leaving examinations (whether English or Scottish) as a proxy for the type of degree. The identifying

assumption for our main regression equation is that students in England and Scotland are no different on unobservable characteristics, i.e. $\text{Cov}(\varepsilon_{ij}, SCOT_{ij}) = 0$.

Although this paper is focused on wage comparisons, I also consider the effect of a Scottish degree on the probability of switching:

$$SWITCH_{ij} = \theta'X_{ij} + \mu SCOT_{ij} + \phi_j + \varepsilon_{ij}, \quad [2]$$

where $SWITCH_{ij}$ and $SCOT_{ij}$ are as defined in equation [1]. The set of controls, X_{ij} , includes sex, age, marital status, high school GPA, and parental socioeconomic status. Some specifications also include field of study effects and controls for region of work. In this regression, λ captures the difference between England and Scotland in the likelihood of switching. Again, I use the type of school leaving examinations to estimate a reduced-form equation of the probability of field switching.

6. Results

6.1 Wages

Wage regressions are presented in Table 2. Columns (1), (2), and (3) explore the effects on wages in the first job held in the first year after completing a BA degree, whereas columns (4), (5), and (6) examine the effects on wages in the job held 6 years after completing a BA degree. In addition to gender, marital status, age, high school GPA, honors level, all wage regressions

Table 2. The effect of Scottish degree and field switching on log annual earnings

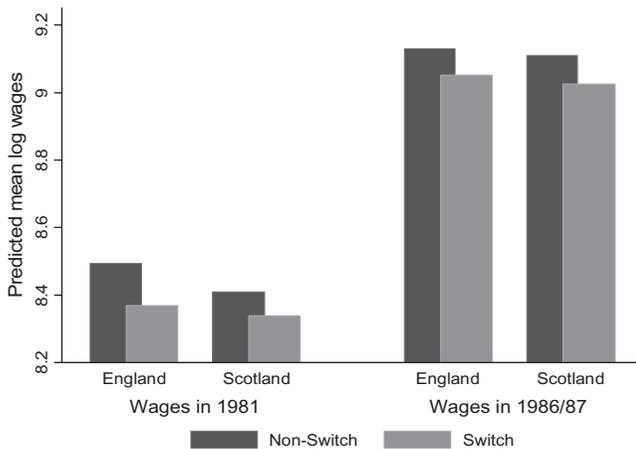
	First year after completing degree			Sixth year after completing degree		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>SCOT</i>	-0.033 [0.035]		-0.057 [0.034]	-0.014 [0.050]		-0.034 [0.053]
<i>SWITCH</i>		-0.061* [0.026]	-0.070** [0.026]		0.003 [0.029]	-0.004 [0.030]
<i>SCOT*SWITCH</i>			0.069 [0.049]			0.057 [0.057]
Main controls	X	X	X	X	X	X
Field of study effects	X	X	X	X	X	X
Region of work effects	X	X	X	X	X	X
R^2	0.24	0.24	0.25	0.31	0.31	0.31
Observations	1,379	1,379	1,379	1,374	1,374	1,374
Mean of dependent variable	8.43	8.43	8.43	9.12	9.12	9.12

Notes: Huber–White standard errors in brackets. * and ** indicate significance at the 5 per cent and 1 per cent level, respectively. Sample includes all students who aimed to attain a first degree in England and Scotland with occupation data and were not pursuing further studies. Dependent variable in columns (1), (2), and (3) is defined as log annual earnings in the first year after completion of an undergraduate degree. Dependent variable in columns (4), (5), and (6) is defined as log annual earnings in the sixth year after completion of an undergraduate degree (deflated to 1981 pounds). *SWITCH* is defined as 1 if the broad field of study is different from the broad occupational field of first job in the first year following degree and 0 otherwise. *SCOT* is defined as 1 for Scottish high school exams and 0 for English high school exams. Main controls include sex, marital status, age, high school GPA, parent SES, and industry fixed effects.

include controls for field of study, industry, and region of work as wages may differ markedly across fields, regions, and industry for other reasons. Column (1) reveals that there is no significant difference in average annual earnings between England and Scotland in the first year following completion of the degree — the coefficient on *SCOT* from equation [1] is not statistically significant. According to the model, this suggests that the benefits to higher match quality from later specialization are roughly in balance with the benefits to specific skills from early specialization. However, this is also consistent with the discussion of Proposition 3 and evidence from Figure 2 that differences in expected wages between England and Scotland tend to narrow when the return to match quality is relatively high (as suggested by the pattern of switching described in the following section).²⁹ Column (2) provides strong evidence supporting the prediction of Proposition 2 that individuals who switch to an occupation unrelated to their field of study at university earn lower wages in the first year (1981) — the coefficient on *SWITCH* is negative and significant. Indeed, field switching is associated with a substantial wage loss of around 7 percentage points, comparable in magnitude to the negative wage differential for women in this sample. The magnitude of the coefficient on *SCOT*SWITCH* in column (3) suggests that the differential associated with field switching is larger in England than Scotland but the estimate is rather imprecise.³⁰

Column (4) shows that there is also no significant difference in average annual earnings between England and Scotland after 6 years in 1986/87. Interestingly, columns (5) and (6) indicate that individuals who switched to an occupation unrelated to their field of study at university in the first year earn average annual wages 6 years later that are no different from their counterparts who did not switch. In other words, controlling for background variables, individuals who experience field switching appear to make up the difference over time. If field-specific skills are also accumulated on the job but these skills have diminishing returns, we would expect the wage loss associated with switching to decline over time. Figure 4 plots log wages in 1981 and 1986/87 predicted on the basis of observable characteristics from the wage regressions of Table 2; specifically, columns (3) and (6). Although insignificant, the differential in initial wages between those who switch and those who do not switch does

Figure 4. Predicted log wages in 1981 and 1986/87



Notes: Mean log wages are predicted based on observable characteristics from the wage regressions of Table 2; specifically, columns (3) and (6). Log wages in 1986/87 are deflated to 1981 prices.

appear to be larger in England than in Scotland. Robustness checks for these findings are presented in columns (2), (3), and (4) of Table A2. In particular, the findings are robust to alternative (broader and narrower) classifications, as well as the sampling restrictions described in the previous section. Moreover, the findings remain unchanged when restricting to the sample of students with top high school GPAs and when excluding students at Oxford and Cambridge.

Part of the wage loss among individuals who switch to unrelated occupations may be associated with unobservables, which are correlated with field switching rather than a direct causal effect. However, it is important to distinguish between two types of unobserved variables. Although we try to control for ability using high school achievement and success in university, individuals may have additional unobservable traits that affect wages. For example, individuals who are particularly indecisive and therefore switch fields, as suggested in the previous section, may have ended up earning lower wages in any case.³¹ In the model of academic specialization presented earlier, switching fields is endogenous yet individuals switch fields because they receive new information on match quality and not because of some unobserved characteristics. Shocks to information on match quality will generally be unobservable.³² But these reflect the inherent uncertainty in the process of learning about match quality rather than an innate trait. The fact that individuals who switch tend to catch up with their counterparts who did not switch suggests that this may be the more important explanation for the initial wage loss associated with switching.

6.2 Field switching

Table 3 examines field switching between England and Scotland. We estimate a reduced-form equation where *SCOT* is a dummy variable identifying whether students took English or Scottish secondary school leaving exams. All regressions include controls for gender, marital

Table 3. Effect of Scottish degree on field switching for 1980 college graduates

	First year after completing degree			Sixth year after completing degree		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>SCOT</i>	-0.151** [0.035]	-0.086** [0.028]	0.014 [0.045]	-0.174** [0.036]	-0.110** [0.029]	-0.053 [0.046]
Main controls	X	X	X	X	X	X
Field of study effects		X	X		X	X
Region of work effects			X			X
R^2	0.03	0.35	0.36	0.03	0.30	0.31
Observations	1,455	1,455	1,455	1,455	1,455	1,455
Mean of dependent variable	0.48	0.48	0.48	0.52	0.52	0.52

Notes: Huber–White standard errors, clustered by university in brackets. * and ** indicate significance at the 5 per cent and 1 per cent level, respectively. Sample includes all students who aimed to attain a first degree in England and Scotland with occupation data and were not pursuing further studies. Dependent variable is defined as 1 if field of study at the undergraduate level is different from the broad occupational field of the first job in the first year following the degree and 0 otherwise. *SCOT* is defined as 1 for having completed Scottish school leaving exams and 0 for English school leaving exams. Main controls include sex, marital status, age, high school GPA, and parent SES.

status, age, socio-economics status (SES) level, and high school GPA. In column (1), I estimate the difference in the probability of field switching between England and Scotland without controlling for fields of study or region of work. Similar to the raw difference shown in Table 1, the estimated difference in field switching is approximately 15 percentage points, which is substantial given that the rate of field switching in Scotland is about 0.34. According to the model, these findings suggest that the return to match quality is high relative to the return to specific skills. Once I control for the composition of fields across nations in column (2), the estimated differential in field switching declines substantially. In other words, not only do individuals in Scotland switch less, but they also tend to study fields that are associated with less switching.³³ In column (3), I add controls for region of work and the coefficient on *SCOT* becomes insignificant, suggesting that there may be different preferences for related qualifications among employers in England and Scotland. However, this specification needs to be interpreted with care as the decision to work in England or Scotland is probably endogenous; individuals who decide to switch may also make systematically different decisions about where they wish to work.

The NSGD also contains information on student outcomes 6 years following the completion of their degree. Columns (4), (5), and (6) indicate that the differential in field switching between England and Scotland remains after 6 years. Even stronger results are obtained if we consider all individuals employed 6 years following completion of the BA degree by including those who were not employed in the first year after completing their degree (results not shown). This suggests that individuals in England continue to experiment more than individuals in Scotland once in the labor market. Robustness checks for all these findings are presented in column (4) of Table A2. As the likelihood of switching fields is substantially higher in England than in Scotland, the model suggests that more students in England suffer a loss in specific skills when trying to correct mistakes in their initial choice of field. Thus, at least during the initial years of labor market experience, a regime with early specialization appears to be associated with more costly switches than one with later specialization.

6.3 Other results

Although outside the scope of the model proper, I also consider several labor market outcomes over time. The preceding section on wages documented that individuals who experience field switching make up the difference in wages over time. In a related set of specifications, columns (1), (2), and (3) of Table 4 show the growth in annual wages over the 6 years following completion of a BA degree. With coefficients on *SWITCH* that are almost significant, columns (2) and (3) suggest that individuals who switch to an unrelated occupation upon entering the labor market experience greater wage growth than their counterparts who do not switch. As explained earlier, if field-specific skills are also accumulated on the job but these skills have diminishing returns, we would expect field switching to be associated with greater wage growth in the early years after graduation. Moreover, although insignificant, the signs on *SCOT* and *SCOT*SWITCH* accord with the general intuition: individuals who switch in Scotland experience lower wage growth than their English counterparts as they have higher levels of specific skills upon entering the labor market; individuals who do not switch in Scotland experience greater wage growth than their English counterparts as they have lower levels of specific skills upon entering the labor market. Columns (4), (5), and (6) of Table 4 explore occupational mobility, which is defined as a further change in occupational field following entry into the labor market.³⁴ Individuals who experience field switching are significantly more likely to change to a job in a different occupational field after several years. This

Table 4. The effect of Scottish degree and field switching on wage growth and occupational mobility

Dependent variable	Growth in log annual earnings			Occupational mobility		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>SCOT</i>	0.034 [0.058]		0.039 [0.054]	-0.096* [0.048]		-0.137** [0.049]
<i>SWITCH</i>		0.065 [0.036]	0.066 [0.037]		0.065* [0.031]	0.052 [0.032]
<i>SCOT*SWITCH</i>			-0.016 [0.074]			0.108 [0.065]
Main controls	X	X	X	X	X	X
Field of study effects	X	X	X	X	X	X
Region of work effects	X	X	X	X	X	X
<i>R</i> ²	0.19	0.19	0.19	0.08	0.08	0.08
Observations	1,379	1,379	1,379	1,455	1,455	1,455
Mean of dependent variable	0.69	0.69	0.69	0.24	0.24	0.24

Notes: Huber–White standard errors in brackets. * and ** indicate significance at the 5 per cent and 1 per cent level, respectively. Sample includes all students who aimed to attain a first degree in England and Scotland with occupation data and were not pursuing further studies. Dependent variable in columns (1), (2), and (3) is defined as growth in log annual earnings in the 6 years following completion of the undergraduate degree. Dependent variable in columns (4), (5), and (6) is defined as 1 if occupational field after 6 years is different from the occupational field after 1 year following completion of an undergraduate degree. *SWITCH* is defined as 1 if broad field of study is different from the broad occupational field in the first year following degree and 0 otherwise. *SCOT* is defined as 1 for Scottish high school exams and 0 for English high school exams. Main controls include sex, marital status, age, high school GPA, parent SES, and industry fixed effects.

is consistent with the theoretical analysis because individuals who switch have the same level of specific skills in alternative fields so that further changes in occupation are not as severely penalized. Nevertheless, it is important to keep in mind that some individuals who switch may simply be inherently less stable workers. Furthermore, among individuals who do not switch, those in Scotland are significantly less likely to change to another occupation in later years.

Finally, respondents in the NSGD were asked: ‘On reflection, how beneficial has your [1980] qualification been to you in’: (i) getting an interesting job; (ii) securing a good income; and (iii) becoming a widely educated person. Based on simple unconditional means, Scottish individuals report higher levels of getting an interesting job than their English counterparts (but these are only marginally significant). Table 5 highlights regression results from these subjective assessments.³⁵ Individuals who switch to an occupation unrelated to their field of study are significantly less likely to consider their qualification beneficial in obtaining an interesting job. Indeed, this effect remains strong even after controlling for wages, subjective assessments of securing a good income, and fields of study. It is possible that individuals who switch consider themselves as having interesting jobs but not as a direct result of their qualification. Nevertheless, this provides some suggestive evidence for the non-pecuniary benefits of entering an occupation related to the field of study at university. Individuals who experience field switching are significantly more likely to report that their qualification contributed to their becoming more widely educated. However, this effect becomes insignificant once controls for field of study are included, suggesting that individuals who consider themselves widely educated were those who selected fields of study with particularly high rates of switching (e.g. humanities and social sciences).

Table 5. Subjective assessments

Dependent variable: ‘How beneficial has your qualification been to you in . . . ’

	‘Getting an interesting job’		‘Becoming an educated person’	
	(1)	(2)	(3)	(4)
<i>SCOT</i>	-0.154 [0.126]	-0.202 [0.128]	0.041 [0.112]	0.044 [0.116]
<i>SWITCH</i>	-0.265* [0.075]	-0.197* [0.083]	0.312** [0.066]	0.082 [0.178]
<i>SCOT*SWITCH</i>	0.29 [0.189]	0.315 [0.189]	-0.534** [0.177]	-0.536** [0.178]
Main controls	X	X	X	X
Field of study effects		X		X
Observations	1,345	1,345	1,345	1,345
Mean of dependent variable	3.24	3.24	2.83	2.83

Notes: Huber–White standard errors in brackets. * and ** indicate significance at the 5 per cent and 1 per cent level, respectively. Sample includes all students who aimed to attain a first degree in England and Scotland with occupation data and were not pursuing further studies. Results are from ordered probit regressions. Dependent variables are ordered categorical variables that take on values from 1 (‘Not at all’) to 4 (‘A lot’). *SWITCH* is defined as 1 if broad field of study is different from the broad occupational field of first job in the first year following degree and 0 otherwise. *SCOT* is defined as 1 for Scottish high school exams and 0 for English high school exams. Main controls include sex, marital status, age, high school GPA, parent SES, industry fixed effects, log annual earnings in the first year, and subjective measures of how beneficial the qualification has been to securing a good income.

7. Conclusion

Specialization is a fundamental feature of many economic decisions. This paper examines the trade-off between the acquisition of specific skills early in one’s college education versus broader training and learning about match quality in different fields of study. I introduce a model of specialization in which individuals accumulate field-specific skills and receive noisy signals of match quality by taking courses in different fields of study. I also extend the model to allow for switching to occupations, which are unrelated to the chosen field of study, upon entry into the labor market. I then derive comparative static predictions between regimes with early and late specialization and test them across the corresponding English and Scottish systems of higher education. I find strong evidence in support of the prediction that individuals who switch to unrelated occupations initially have, on average, lower wages. Although not significant, the magnitude of the wage differential associated with switching suggests a larger wage penalty in England than Scotland. Moreover, as the probability of switching fields is substantially higher in England than in Scotland, more individuals in England suffer a loss in initial wages. Nevertheless, I also find that individuals who switch tend to experience greater wage growth so that wage differences between those who do and do not switch become insignificant after 6 years in the labor market. Furthermore, controlling for demographic and occupational characteristics, there is no significant difference in average wages or subjective satisfaction between England and Scotland.³⁶ Together, these findings suggest that, although early specialization is associated with more costly switches upon entering the labor market, differences in wages across educational systems with early and late specialization do not persist over time.

How should we interpret the absence of significant differences in average earnings between England and Scotland? Based on the model, this result suggests that the benefits to higher match quality from later specialization is roughly in balance with the benefits to more specific skills from early specialization. On the other hand, the pattern of switching indicates that the return to match quality is relatively high as compared with the return to specific skills. There are several ways to reconcile these findings. First, it is possible that the benefits of later specialization are mostly associated with non-pecuniary considerations, which are captured by field switching but missed in a measure of wages.³⁷ Second, as explained in the discussion of Proposition 3 and shown in Figure 2, differences between early and late specialization tend to diminish when switching is possible and the return to match quality is relatively high, because individuals in the early regime are able to correct for their initial mistakes. Finally, there is a range of parameter values for the relative return to match quality where both regimes yield similar returns even though the early regime has a higher rate of switching.³⁸ Distinguishing between these alternative explanations is an important area for further research.

Whether early or late specialization is optimal also depends on other structural parameters, such as the accuracy of information on match quality prior to commencing education. Lower prior variances on match quality to different fields would imply that a shorter period of experimentation is optimal. Hence, if certain populations have more accurate information about their match quality, a regime with early specialization may not necessarily lead to many costly mistakes. Indeed, with a heterogeneous population, the benefits associated with early or late specialization may accrue to different individuals. The theoretical model abstracts from distributional concerns by assuming that individuals are identical but these considerations may, in fact, be important in evaluating different systems of higher education. In a broader sense, however, both distributional considerations and possible inefficiencies arise because individuals are forced to specialize at a particular time. Allowing individuals to choose when to specialize should be optimal, especially with a heterogeneous population.

The American system of higher education is often cited as an example of an educational system with a broad undergraduate curriculum and relatively late specialization. However, although the USA has a strong liberal arts tradition that emphasizes general education and allows for later specialization, American undergraduates can often exploit their elective courses and specialize early, if they wish. The US system of higher education is not so much characterized by a broad curriculum and late specialization as by flexibility in breadth and timing of academic specialization. As mentioned above, a system that allows individuals to choose when to specialize should be optimal if students are able to experiment in a productive fashion. But if there is some lumpiness in human capital investments and students do not know how to experiment in a productive manner, such flexibility can lead students to end up with unproductive programs of study. For example, Trow (1999) has argued that American undergraduates often take incoherent courses of study and indulge in excessive experimentation. Moreover, individuals who choose to specialize later may also take longer to graduate from college and therefore generate additional social costs. Examining whether the flexibility of the American educational system can improve on both the English and Scottish systems would be a valuable next step in extending our knowledge of academic specialization.

Appendix A

Complete documentation for the National Survey of 1980 Graduates and Diplomates, 1986–87 is available from the UK Data Archive: <http://www.data-archive.ac.uk> (Department

of Employment, 1988). Details of the variables constructed for this study are described as follows.

Field switch

A field switch is defined as a binary variable that takes on a value of 1 if an individual is employed in an occupation that is unrelated to his or her major field of study at the undergraduate level, and 0 otherwise. In order to determine whether an individual is employed in an occupation that is related or unrelated to his or her field of study, I group fields of study and occupations into categories. As shown in Table A1, I allow for three gradations of classification: narrow (42 categories), broad (12 categories), and very broad (6 categories). Occupations and fields of study are coded according to each of the alternative classifications. Where the occupation and field of study are classified in different categories, the field switch variable takes on a value of 1. For example, an individual who studies physics at university will have his or her field of study coded as 'physics' according to the narrow classification, 'physical sciences' according to the broad classification, and 'mathematical, computer, and physical sciences' according to the very broad classification. If this individual is employed as a computer programmer, the field switch variable will take on a value of 1 according to the narrow and broad classifications and a value of 0 according to the very broad classification. Combined fields are considered switches if the individual is not employed in any of the fields mentioned. I focus on the broad classification in most of the analysis in this paper. I also consider field switching from the graduate level. This variable is defined analogously except that the field of study is the one studied at the graduate level.

High school GPA

Scores on secondary school leaving exams are officially coded as letter grades (A, B, C, etc.). These are converted into numerical scores where A = 10, B = 8, C = 6, D = 4, and E = 2. Average scores are then standardized by nation and combined so that the overall distribution of high school GPA has mean 0 and standard deviation 1.

SES

Individual SES scores are based on parental occupations as follows: 0 — unstated, retired, or unknown, 1 — professionals workers, 2 — intermediate workers, 3 — skilled non-manual, 4 — skilled manual, 5 — partially skilled, 6 — unskilled, and 7 — unemployed.

Region of work

Region of work is classified as London, Southern England, Midlands, East Anglia, Northern England, Wales, Scotland, Northern Ireland, or abroad in the NSGD data.

Industry

Industry are classified according to broad Standard Industry Codes: Agriculture, Forestry, and Fisheries (0), Mining (1), Mineral Extraction and Production (2), Heavy Manufacturing

Table A1. Classification of fields and occupations

Fields	Subject codes (NSGD/USR-1980)	Occupational codes (NSGD)	Occupational codes (USR-1980)
111 Math/Comp. Science			
1111 Math Sciences	Mathematics (81)	Mathematician (444); Statistician (242) . . .	Operational research (441); Statistician (452)
1112 Computer Sciences	Computer Science (82); Math/Comp. Science (31)	Computer Programmer (244); Analyst/programmer (246) . . .	Systems analysis (442); Computer programming (443) . . .
112 Physical Sciences			
1121 Chemistry	Chemistry (34); Environmental Science (36)	Chemical scientist (442)	Scientist (510) + Chemical and allied industries (240–247)
1122 Geology	Geology (35)	Geological scientist (445)	Scientist (510) + Oil, mining industries (230–235)
1123 Physics	Physics (33); Mathematics/Physics (32)	Physical scientist (443)	Scientist (510) + Atomic energy (284); Other manufacturing
121 Architecture	Architecture (51); Town plan (52); Surveying (17)	Architect (511); Town planning (514); Draughtsman (490) . . .	Architect (551); Town planning (553); Surveying (554) . . .
122 Engineering			
1221 Mechanical	Mechanical engineering (12)	Mechanical or aeronautical engineer (461)	Engineer (520) + Automotive industry (253)
1222 Chemical	Chemical Engineering (9)	Chemical engineer (481)	Engineer (520) + Chemical and allied industries (240–247)
1223 Civil	Civil Engineering (10)	Civil, municipal or structural engineer (451) . . .	Engineer (520) + Civil engineering contractors (220–225) . . .
1224 Electrical	Electrical Engineering (11)	Electrical engineer (471); Electronic engineer (472/473)	Engineer (520) + Electric (256); Computers (257) . . .
1225 Industrial	Production engineering (13)	Production engineer (482); Planning engineer (483) . . .	Engineer (520) + Food (261); Drink (262); Textiles (271) . . .
1226 Materials	Mining (14); Metallurgy (15)	Mining engineer (452); Metallurgist (485)	Engineer (520) + Oil, mining industries (230–235)
1227 Aeronautical	Aeronautical engineer (8)	Mechanical or aeronautical engineer (461)	Engineer (520) + Aircraft, aerospace industry (254)
131 Life Sciences			
1311 Agriculture	Agriculture (20); Forestry (23) . . .	Farmer, farm manager, horticulturist (600)	Scientist (510) + Agriculture, horticulture, forestry (210–214)
1312 Biology	Biology (25); Botany (26); Zoology (27) . . .	Biological scientist, biochemist (441)	Scientist (510) + Health authorities (154)
132 Health Sciences			
1321 Physicians	Medicine (3)	Medical practitioner (351)	Medicine (631); Medical & para-medical services (630)
1322 Dentists/Vets/Pharm	Dentistry (4); Veterinary (24); Pharmacology (5, 6)	Dentist (352); Veterinarian (382); Pharmacist (371) . . .	Dentistry (632); Veterinary (640); Pharmacy (634) . . .
1323 Nursing/Related	Studies allied to medicine/health (7)	Nurse (360); Physiotherapist (374) . . .	Nursing (633); Physio-occupational, speech & therapy (636)
211 Social Service			
2111 Psychology	Psychology (46)	Psychologist (324)	Psychology (623); Occupational guidance (624)
2112 Sociology/Social Work	Sociology (47)	Sociologist (323); Welfare worker (333) . . .	Social, welfare, religious (620); Social/welfare (621) . . .
212 Social Sciences			
2121 Economics	Economics (41)	Economist (241)	Economic (450); Economist (451)
2122 History/Geography	History (69); Archeology (70); Geography (42)	Librarian, information officer (294)	Librarian (721); Archivist (722)
2123 Govt., Public Admin.	Government and public administration (44)	Inspector (263); General administration (local govt) (280) . . .	Consumer protection, environmental health, safety (653) . . .
2124 Other Social	Social anthropology (48)	Social or behavioral scientist (325)	Non-scientific research (730); Information research (700) . . .
221 Business			
2211 Accounting, Finance	Accountancy (43)	Accountant (221); Investment analyst (228) . . .	Financial (460); Accountancy (461); Banking (462) . . .
2212 Management	Business, management studies (40, 53)	Management consultant (296); Manager (561) . . .	Management & supporting occupations (400) . . .
2213 Sales	Business, management studies (40, 53)	Advertising executive (252); Buying and selling (255) . . .	Purchasing (431); Selling (432); Marketing (434) . . .
2214 Related Business	Secretarial studies (84)	Office manager (572); Personal assistant (297) . . .	Clerical, secretarial & related (930) . . .
222 Law	Law (45)	Judge (211); Advocate, barrister (212); Solicitor (213) . . .	Baristor (471); Solicitor (472); Trusts (473) . . .
231 Education	Education (1)	Teacher (secondary) (311); Teacher (primary) (312) . . .	Primary (611); Middle school (612); Secondary (613) . . .
232 Arts			
2321 English/Languages	English (65); French (57); German (59) . . .	Author, writer, journalist, editor (391)	Journalist (811); Technical writer (711); Translator (712) . . .
2322 Art	Art (73)	Artist, commercial artist (401); Designer (402–406)	Art, sculpture, design (820); Fashion & textiles (823) . . .
2323 Performing arts	Drama (74); Music (75)	Actor, entertainer, musician, singer, stage manager (411) . . .	Acting, music, sport (830); Broadcasting/stage/film (840) . . .
2324 Religion/Philosophy	Religion (72); Philosophy (71)	Clergy, minister of religion (340)	Pastoral (622)

Notes: Subject codes for USR are correct for 1972–84 (different codes for 1985–93) and occupational codes for the USR are correct for 1980–93 (different codes for 1973–79). Occupational codes omit some categories for brevity and indicated with . . . when excluded. Engineers and scientist in the USR are matched with industry codes in order to identify particular specializations within each category. Further details are available from the author. Broad fields are in bold. Very broad fields are expressed by the two-digit codes.

Table A2. Robustness checks

Coefficient of interest	Differential in wages (<i>SCOT</i>) (1)	Wage loss (<i>SWITCH</i>) (2)	Differential in wage loss (<i>SCOT</i> * <i>SWITCH</i>) (3)	Differential in switching (<i>SCOT</i>) (4)
Baseline	-0.057 [0.034]	-0.070** [0.026]	0.069 [0.049]	-0.086** [0.028]
Classification of fields				
Narrow	-0.089* [0.038]	-0.067** [0.021]	0.102* [0.042]	-0.062 [0.034]
Very broad	-0.028 [0.039]	-0.071* [0.029]	-0.023 [0.042]	-0.070* [0.028]
Occupational restrictions				
Unemployed as switch	-0.057 [0.034]	-0.070** [0.027]	0.069 [0.049]	-0.089** [0.028]
Unclassified occupations as switch	-0.054 [0.033]	-0.065** [0.022]	0.026 [0.044]	-0.064* [0.026]
Include graduate students	-0.055 [0.034]	-0.069** [0.026]	0.074 [0.049]	-0.088** [0.028]
Graduate students as non-switch	-0.057 [0.034]	-0.070** [0.026]	0.069 [0.049]	-0.086** [0.028]
Field restrictions				
Exclude education	-0.062 [0.034]	-0.072** [0.027]	0.079 [0.050]	-0.092** [0.029]
Education as non-switch	-0.061 [0.034]	-0.071** [0.027]	0.079 [0.049]	-0.089** [0.028]
Exclude business	-0.055 [0.037]	-0.080** [0.027]	0.069 [0.050]	-0.079** [0.030]
Business as non-switch	-0.058 [0.034]	-0.080** [0.027]	0.074 [0.049]	-0.074** [0.028]
Exclude combined fields	-0.026 [0.032]	-0.078* [0.032]	0.062 [0.051]	-0.107** [0.028]
Population restrictions				
Restrict to highest GPA students	-0.138** [0.050]	-0.071 [0.040]	0.089 [0.106]	-0.081 [0.065]

Notes: Huber-White standard errors in brackets. *, ** indicate significance at the 5 per cent and 1 per cent level, respectively. Coefficients on all wage regressions [columns (1)-(3)] include main controls, controls for field of study, region of work, and industry. Coefficients on occupational switching sample [column (4)] include main controls, controls for field of study, but no controls for region of work.

(3), Light Manufacturing (4), Construction (5), Wholesale and Retail Trade (6), Transportation, Communication, and Public Utilities (7), Financial and Business Services (8), Professional and Related Services (9).

Wages

Wages are annual earnings as reported by individuals in a retrospective survey. Accurate measures of wages are available at two times: starting wages in the first job and ‘current wages’ in the last job reported approximately 6 years after completing the first degree.

Appendix B

This mathematical appendix provides a formal treatment of the model of academic specialization presented in the main text. For ease of exposition, the structure of the appendix and most of the notation parallels the main text.

Formal set-up

Suppose N courses are taken in $k \geq 2$ fields of study. Let F_1, \dots, F_k be normal populations associated with fields of study $i = 1, \dots, k$, each with unknown mean $\theta_1, \dots, \theta_k$ and a common known variance $\sigma^2 > 0$. The unknown means $\theta_1, \dots, \theta_k$ represent unobserved match quality in each field.

Sequence of observations. In Stage 1, n observations from each population F_i are observed. These correspond to observations on match quality from courses taken in each field of study prior to specialization. The sample means of these observations, X_i , are independent and distributed $N(\theta_i, p^{-1})$ with $p = n\sigma^{-2}$. In Stage 2, one population, i^* , is selected for further sampling and $(N - nk)$ additional observations are observed from this population. These correspond to observations on match quality in the chosen field from courses taken following specialization. The sample mean of the second set of observations, Y , is distributed $N(\theta_{i^*}, q^{-1})$ with $q = (N - nk)\sigma^{-2}$ and where θ_{i^*} is the (unknown) mean of the population chosen after Stage 1.³⁹

Beliefs on match quality. Belief about match quality $\theta_1, \dots, \theta_k$ are represented by the parameters $\hat{\theta}_1, \dots, \hat{\theta}_k$. These parameters are random and follow independent and identical prior distributions assumed to have $\hat{\theta}_i \sim N(\mu, \nu^{-1})$ with $\nu = \sigma_0^{-2}$. The conditional distribution of $\hat{\theta}$ at each stage can be expressed as follows:

$$\hat{\theta}_i | \mathbf{X} = \mathbf{x} \sim N\left(\frac{px_i + \nu\mu}{p + \nu}, (p + \nu)^{-1}\right), \quad i = 1, \dots, k \text{ independent}$$

$$\hat{\theta}_i | \mathbf{X} = \mathbf{x}, Y = y \sim N\left(\frac{\pi\mu_i(\mathbf{x}) + q_i Y}{\pi + q_i}, (\pi + q_i)^{-1}\right), \quad q_{i^*} = q \text{ and } 0 \text{ otherwise,}$$

where $\pi = p + \nu$ represents the relative combined (prior plus sampling) information gained from field F_i , and where $\mu_i(\mathbf{x}) = (px_i + \nu\mu)/(p + \nu)$ represents the estimated mean of field F_i after Stage 1. In terms of the notation in the main text, $\mu'_i = \mu_i(\mathbf{x})$ and $\mu''_i = \mu_i(\mathbf{x}, y)$.⁴⁰

Pay-offs. The pay-off associated with field F_i is denoted by $w_i = \alpha\theta_i + \beta s_i$ where s_i is the cumulative number of observations from field F_i . This pay-off represents the wage received in field i upon entering the labor market. In terms of the model of academic specialization, α is the return to match quality and β is the return to specific skills. Note that we can express the loss function associated with population F_i as $L_i(\theta, s) = \alpha\theta_i - \beta s_i$.⁴¹

Decision rules

After $\mathbf{X} = \mathbf{x}$ has been observed at Stage 1, the Bayes selection rule $i^* = d_1^*(\mathbf{x})$ can be found by minimizing the posterior expected loss (or in our framework, maximizing posterior expected wages):

$$\begin{aligned} E_X\left(L\left(\hat{\theta}, d_1^*(\mathbf{X})\right) \mid \mathbf{X} = \mathbf{x}\right) &= \max_{i=1, \dots, k} E_X\left(\alpha\hat{\theta}_i + \beta s_i \mid \mathbf{X} = \mathbf{x}\right) \\ &= \alpha \max_{i=1, \dots, k} E_X\left(\hat{\theta}_i \mid \mathbf{X} = \mathbf{x}\right) + \beta s \\ &= \alpha \max_{i=1, \dots, k} \mu_i(\mathbf{x}) + \beta s = \alpha \left(\frac{p(\max_{i=1, \dots, k} x_i) + v\mu}{p + v} \right) + \beta s, \end{aligned}$$

where s corresponds to the specific skills in each field, which are equivalent across fields. The optimal selection, i^* , at Stage 1 will therefore be the population with the largest observed sample mean after Stage 1 as $d_1^*(\mathbf{x}) = \arg \max_{i=1, \dots, k} x_i$. This is intuitive as, with identical prior distributions on match quality, the only distinguishing feature of each population is the information received in Stage 1. Let $x_{[1]} < x_{[2]} < \dots < x_{[k]}$ denote the order sample means from Stage 1 and $\mu_{[1]}(\mathbf{x}) < \mu_{[2]}(\mathbf{x}) < \dots < \mu_{[k]}(\mathbf{x})$ denote the ordered posterior means from Stage 1. Note that, in terms of the notation in the main text, $\mu'_{i^*} = \mu_{[k]}(\mathbf{x})$ and $\mu'_{i^*} = \mu_{[k-1]}(\mathbf{x})$.

Similarly, after $Y = y$ has been observed at Stage 2, the Bayes selection rule $i^{**} = d_2^*(\mathbf{x}, y)$ will satisfy

$$E\left(L\left(\hat{\theta}, d_2^*(\mathbf{X}, \mathbf{Y})\right) \mid \mathbf{X} = \mathbf{x}, Y = y\right) = \max_{i=1, \dots, k} E\left(\alpha\hat{\theta}_i + \beta s_i \mid \mathbf{X} = \mathbf{x}, Y = y\right).$$

These Bayes selection rules yield the maximum posterior expected wages, or Bayes risk, of their respective problems in Stages 1 and 2. Let $\mu''_{i^*} = \mu_{[k]}(\mathbf{x}, y)$ denote the posterior mean of field, i^* , after Stage 2. An important feature of this decision problem is that the selection $i^{**} = d_2^*(\mathbf{x}, y)$ after Stage 2 may differ from the selection $i^* = d_1^*(\mathbf{x})$ after Stage 1 as further observations in Stage 2 may reveal that the initial choice was not as good as initially thought. This corresponds precisely to the possibility of switching fields expressed in the main theoretical framework.

Proof of Proposition 1

Consider the case where switching after Stage 2 is not possible. In this case, the optimal choice of the number of observations, n , sampled from each population in Stage 1, on the Bayes risk associated with the selection in Stage 1 can be determined by maximizing the following expression with respect to n :

$$\begin{aligned}
 E[w(\theta_{i^*}, s)] &= E_X \left[\max_{i=1, \dots, k} E(\alpha \hat{\theta}_i + \beta s_i | \mathbf{X} = \mathbf{x}) \right] \\
 &= \alpha E[\mu_{[k]}(\mathbf{x})] + \beta [N - n(k - 1)] \\
 &= \alpha(\mu + \eta E[Z^{[k]}]) - n\beta(k - 1) - \beta N
 \end{aligned}$$

where $\eta^2 = \frac{p}{v(p+v)} = \frac{n}{nv + \sigma^2 v^2}$ and $Z^{[k]} \sim \max_{i=1, \dots, k} Z_i$ where $Z_i \sim N(0, 1)$.

The last equality follows from a result in order statistics: essentially, we express the maximum of $\mu_i(\mathbf{X})$ in terms of its z -score, the maximum of standard normal distributions Z_i (see Afonja, 1972; Dunnet, 1960). η^2 is the variance of the marginal distribution of $\mu_i(\mathbf{x})$.⁴² Intuitively, $E[\mu_{[k]}(\mathbf{x})]$ represents the return to match quality whereas $[N - n(k - 1)]$ represents the return from specific skills in choosing field i^* . Taking the derivative of the above expression with respect to n , setting it equal to zero, and simplifying yields the optimal number of observations, n^* , to be sampled in Stage 1:

$$n^*(n^* + v\sigma^2)^3 = v\sigma^2 \left(\frac{\alpha\sigma E[Z^{[k]}]}{2\beta(k-1)} \right)^2.$$

The unique positive root of the equation above represents the optimal number of observations, n^* .⁴³ The optimal n is clearly increasing in α/β and decreasing in k . In other words, a regime with late specialization, n^L , will yield higher wages than a regime with early specialization, n^E , if the return to match quality is sufficiently higher than the return to specific skills. ■

Proof of Corollary 1

We can express this directly as follows:

$$E[w^L(\theta, s)] > E[w^E(\theta, s)] \Leftrightarrow \frac{\alpha}{\beta} > \Omega = \frac{(n^L - n^E)(k - 1)}{E[Z^{[k]}](\eta^L - \eta^E)} > 0,$$

where η^L and η^E are the η for the regime with late and early specialization, respectively. So a regime with late specialization is preferred if the returns to match quality are large relative to returns to specific skills. ■

Proof of Proposition 2

Consider the case where all switches are exogenous: i.e. some individuals are forced to switch without regard to the signals they receive in the Stage 2. In this case, the difference in expected wages between those individuals who switch and those who do not switch will simply be $E[w(\theta_{i^*}, s)] - E[w(\theta_{i^a}, s)]$. Following the derivation in the proof of Proposition 1, this expression can be shown to be strictly positive:

$$\begin{aligned}
 E[w(\theta_{i^*}, s)] - E[w(\theta_{i^a}, s)] &= \{\alpha E[\mu_{[k]}(\mathbf{x})] + \beta [N - n(k - 1)]\} - \{\alpha E[\mu_{[k-1]}(\mathbf{x})] + \beta n\} \\
 &= \alpha(\mu + \eta E[Z^{[k]}]) - \alpha(\mu + \eta E[Z^{[k-1]}]) + \beta(N - kn) \\
 &= \alpha\eta E[Z^{[k]} - Z^{[k-1]}] + \beta(N - kn) > 0
 \end{aligned}$$

as $E[Z^{[k]} - Z^{[k-1]}] > 0$ where $Z^{[k]} \sim \max_i Z_i$, $Z^{[k-1]} \sim \max_{i \neq i^*} Z_i$ and $Z_i \sim N(0, 1)$.

In the actual model, however, switching is endogenous. Consequently, expected wages have to be evaluated conditional on beliefs about the wages expected from switching and from not switching: $E[w(\theta_{i^*}, s)|w(\mu''_{i^*}, s) > w(\mu'_{i^*}, s)] - E[w(\theta_{i^*}, s)|w(\mu'_{i^*}, s) > w(\mu''_{i^*}, s)]$. These can be expressed as follows:

$$E[w(\theta_{i^*}, s)|w(\mu''_{i^*}, s) > w(\mu'_{i^*}, s)] = \alpha E[\mu_{[k]}(\mathbf{x})|\mu_{[k]}(\mathbf{x}) > \underline{\lambda}] + \beta[N - n(k - 1)]$$

$$E[w(\theta_{i^*}, s)|w(\mu'_{i^*}, s) > w(\mu''_{i^*}, s)] = \alpha E[\mu_{[k-1]}(\mathbf{x})|\mu_{[k-1]}(\mathbf{x}) < \bar{\lambda}] + \beta n$$

$$\underline{\lambda} = \mu_{[k-1]}(\mathbf{x}) - \left(\frac{q}{\pi(\pi + q)}\right)^{1/2} Y^{[k]} - \frac{\beta}{\alpha}(N - nk), \quad \bar{\lambda} = \mu_{[k]}(\mathbf{x}) + \left(\frac{q}{\pi(\pi + q)}\right)^{1/2} Y^{[k]} + \frac{\beta}{\alpha}(N - nk).$$

As these expressions are essentially the distributions associated with exogenous switching but truncated from above and below, it is clear that $E[w(\theta_{i^*}, s)|w(\mu''_{i^*}, s) > w(\mu'_{i^*}, s)] > E[w(\theta_{i^*}, s)]$ and that $E[w(\theta_{i^*}, s)|w(\mu'_{i^*}, s) > w(\mu''_{i^*}, s)] < E[w(\theta_{i^*}, s)]$. Hence, the difference remains positive in the case of endogenous switching. ■

Remarks on Proposition 3

In the case where switching after Stage 2 is possible, determining the optimal number of observations, n , sampled from each population in Stage 1 requires us to evaluate $E[w(\theta_{i^*}, s)]$. Using the standard notation from earlier proofs, we can rewrite this expression in terms of the expectation of the maximum between the chosen field, i^* , and the second best field, i' . We can further decompose $\mu_{[k]}(x, y)$, into the mean of match quality in the chosen field after Stage 1, $\mu_{[k]}(x)$, and the mean of the observations taken from the chosen field in Stage 2, $Y^{[k]}$.⁴⁴

$$\begin{aligned} E(\max[w(\mu''_{i^*}, s), w(\mu'_{i^*}, s)]) &= E_{X,Y} \left(\max \left[\alpha \frac{\pi \mu_{[k]}(\mathbf{X}) + q Y^{[k]}}{\pi + q} + \beta s_{[k]}, \alpha \mu_{[k-1]}(\mathbf{X}) + \beta s_{[k-1]} \right] \right) \\ &= E \left(\max \left[\alpha \mu_{[k]}(\mathbf{X}) + \alpha \frac{q}{\pi + q} [Y - \mu_{[k]}(\mathbf{X})] + \beta [N - n(k - 1)], \alpha \mu_{[k-1]}(\mathbf{X}) + \beta n \right] \right) \\ &= \alpha E[\mu_{[k]}(\mathbf{X})] + \beta [N - n(k - 1)] + E \left(\max \left[\frac{\alpha q}{\pi + q} [Y^{[k]} - \mu_{[k]}(\mathbf{X})], \Delta_{[k]}(\mathbf{X}) \right] \right) \end{aligned}$$

where $\Delta_{[k]}(\mathbf{X}) = \alpha(\mu_{[k-1]}(\mathbf{X}) - \mu_{[k]}(\mathbf{X})) + \beta(nk - N) < 0$.

The first two terms represent the returns to match quality and specific skills as in the case where no switching is permitted. The impact of switching is captured by the final term, which is non-negative in expectation as Y is distributed with mean $\mu_{[k]}(\mathbf{X})$. Clearly, the possibility of switching can only serve to increase expected wages. Note that $\Delta_{[k]}(\mathbf{x}) < 0$ as switching will lead to a loss in specific skills $\beta(nk - N) < 0$ and a loss in match quality $\mu_{[k-1]}(\mathbf{X}) - \mu_{[k]}(\mathbf{X}) < 0$.

The effect of n on expected wages when switching is possible depends on the derivative of the term $E \left(\max \left[\frac{\alpha q}{\pi + q} [Y^{[k]} - \mu_{[k]}(\mathbf{X})], \Delta_{[k]}(\mathbf{X}) \right] \right)$ with respect to n . Unfortunately, it is not possible to evaluate it analytically. However, we can see that the term Ψ will diminish in importance when β is large relative to α . Intuitively, if the returns to specific skills are large, then we do not expect much switching to take place in any case. When α is large relative to β , the derivative of $\mu_{[k]}(\mathbf{X})$ with respect to n will be negative. Then Ψ will be decreasing with n and

provide a countervailing weight to the first term $\alpha E[\mu_{[k]}(\mathbf{X})]$, which is increasing with n . Therefore, allowing for the possibility of switching implies that the highest benefit to switching will occur when n is small. In other words, the possibility of switching is more valuable in a regime with early specialization than one with later specialization. Again, this is intuitive given that more mistakes are made with early specialization and it is therefore more valuable to be able to correct them through switching. The question remains whether Ψ might be sufficiently decreasing in n to overwhelm $\alpha E[\mu_{[k]}(\mathbf{X})]$ and lead a regime with early specialization to dominate the one with late specialization for all α/β . Simulations suggest that this possibility is unlikely. ■

Proof of Proposition 4

Consider again the case where all switches are exogenous. In this case, the effect of the number of observations on match quality prior to specialization, n , on average wage differential from switch clearly depends on α/β :

$$\begin{aligned} \frac{d}{dn} \{E[w(\theta_{i^*}, s)] - E[w(\theta_{i^a}, s)]\} &= \frac{d}{dn} \{\alpha \eta E[Z^{[k]} - Z^{[k-1]}] + \beta(N - kn)\} \\ &= \alpha E[Z^{[k]} - Z^{[k-1]}] \frac{d\eta}{dn} + \frac{d}{dn} \beta(N - kn) \\ &= \alpha \frac{\sigma^2 v^2 E[Z^{[k]} - Z^{[k-1]}]}{2\sqrt{n}(nv + \sigma^2 v^2)^{3/2}} - \beta k \geq 0. \end{aligned}$$

Thus, when the return to match quality is sufficiently large relative to the return to specific skills, the difference in expected wages between non-switchers and switchers will be larger in a regime with late specialization than in one with early specialization, and vice versa. Unfortunately, it is not possible to derive a closed form for effect of n on the expected wage difference between those who do not switch and those who switch. ■

Notes

¹ In a related paper, Malamud (2011), I exploit a similar set-up to test whether higher education, in addition to providing skills, also provides information about one's tastes and talents for different fields of study. The main outcome in that paper is whether an individual works in an occupational field unrelated to their chosen field of study.

² More recently, many English institutions have begun to introduce course structures that include more breadth and offer greater flexibility. This suggests a growing perception that specializing too early may have some drawbacks.

³ This is consistent with evidence from the literature on 'job mismatch' showing that individuals who are over-educated relative to their occupations or under-educated relative to their coworkers earn lower wages. See Sicherman (1991) and Cohn and Kahn (1995) for the USA; Dolton and Vignoles (2000) and McMillen *et al.* (2007) for the UK.

⁴ Relatedly, Nelson and Phelps (1966) and Welch (1970) argue that general education is more valuable than specific training in changing environments because it is more useful for learning new skills and implementing new technologies.

⁵ Learning about match quality is a more prominent feature in models of job turnover. McCall (1990), Miller (1984), Neal (1999), and Shaw (1987) extend the notion of job match quality presented by Johnson (1978) and Jovanovic (1979a) to the occupational level and present evidence for learning about occupational match quality.

⁶ Berger (1988), Grogger and Eide (1995), Hamermesh and Donald (2006), and Rumberger and Thomas (1993) provide evidence that earnings differ by undergraduate major.

⁷ The posterior mean is a weighted average of the prior mean and the mean of the signals: $\mu'_i = (\mu\sigma_0^2 + \sigma^{-2}n\bar{x}_i) / (\sigma_0^2 + n\sigma^{-2})$ where $\bar{x}_i = (1/n)\sum^j x_{ij}$. The posterior variance is $\sigma' = (\sigma_0^2 + n\sigma^{-2})^{-1}$. See DeGroot (1970) for a detailed exposition.

⁸ Strictly speaking, expected future utility should include expected skills rather than the quantity of skills at the point of specialization. But as expected match quality and skills are separable and individuals are risk neutral, this will lead to the same choice at the point of specialization.

⁹ Specifically, $\mu'_i = (\mu\sigma_0^2 + \sigma^{-2}n\max_i \bar{x}_i) / (\sigma_0^2 + n\sigma^{-2})$.

¹⁰ All simulations are based on 5,000 repetitions for $k = 2$, $N = 21$, $\mu_1 = \mu_2 = 0$, $\sigma^2 = 100$, and $\sigma_0^2 = 25$. Early regimes are characterized by $n^E = 2$; late regimes are characterized by $n^L = 6$. Expected wages are determined according to $E(w_i) = E(\alpha\theta_i + \beta\hat{s}_i)$ where $\hat{s}_i = s_i / (N/k) + \mu$ are normalized skills.

¹¹ Indeed, without any learning about match quality, we would need to allow for exogenous switching to generate any switching. Although it is certainly plausible that people switch fields for exogenous reasons, it is not clear why such exogenous switching would differ across educational regimes that feed into what is essentially the same labor market.

¹² This result is expressed formally and proved in Malamud (2011).

¹³ Note, Miller (1984) models job matching as a multi-armed bandit process and derives predictions on the optimal order of sampling jobs. The model of academic specialization in this paper is restricted to a two-stage selection procedure but allows for the simultaneous sampling of different fields.

¹⁴ Of course, the following proposition describes the relationship in the cross-section and not for counterfactual comparison by individuals. As individuals decide optimally, those who decide to switch do better, in expectation, than they would have by remaining in their chosen fields.

¹⁵ Unfortunately, I cannot derive a closed-form solution for the expected wage difference between those who switch and do not switch. For the parameter values used in the simulation (μ_i , σ_0^2 , σ^2 , n^E , n^L , N , and k), the wage differential in the early regime always exceeds the wage differential in the late regime.

¹⁶ There are exceptions: for example, students at Cambridge University are accepted to a broad engineering faculty; students at Keele University are first accepted to complete a year of 'foundation studies'.

¹⁷ Again, there are exceptions: Cambridge's system of Tripos allows some flexibility in making changes to courses of study; the newer universities of Essex, Kent, and Lancaster allow students to study a broader range of subjects.

¹⁸ For example, faculties at the University of Glasgow include Arts, Biomedical and Life Sciences, Education, Engineering, Information and Mathematical Sciences, Law, Business and Social Sciences, Medicine, and Physical Sciences.

¹⁹ Note that changing fields is not always possible. Certain professional faculties, such as medicine and law, are more insular. Engineering is usually a separate faculty but changes from the physical sciences are often permitted.

²⁰ Numerous scholars of British educational systems have noted that Scottish institutions allow for later specialization than English ones: e.g. Evans (1975), Hunter (1971), Osborne (1967), Squires (1987).

²¹ The main exceptions arise in the 'lass' universities established during the 1960s such as the University of Keele, which implemented an experimental modular curriculum.

²² Interestingly, the introduction of A-levels in 1951 to replace the Higher School Certificates was a response to the criticism that these latter qualifications were denying opportunity to pupils with talent in individual subjects who were less successful in others (especially in foreign language requirements). Indeed, the Higher School Certificates had attempted to ensure that pupils followed a sufficiently broad and balanced curriculum by requiring candidates to achieve the minimum standard in a range of subjects for a pass. Dolton and Vignoles (2002) examine the effect of choosing a broader set of courses in secondary school in the UK.

²³ These Scottish qualifications evolved directly from the earlier Leaving and Intermediate Certificates, which required proficiency over a group of subjects rather than in single subjects.

²⁴ This represents a random sample of one in six university graduates. The survey also included graduates from polytechnic and other institutions but I exclude them from the present analysis. Engineering students in Scottish universities are oversampled in the NSGD so it is important to control for fields of study with the NSGD sample.

²⁵ Although there is some choice available with the type of secondary school, through choice of boarding schools, it is undoubtedly much less than in university (the correlation between Scottish residence and attendance in Scottish high school is 0.96). Furthermore, few secondary schools in Scotland offer English leaving examinations (the correlation between attendance in a Scottish high school and sitting Scottish leaving examinations is 0.98). Correlations are derived from the Universities Statistical Record (USR) data.

²⁶ Out of the original sample of 4,889, I lose 510 respondents who cannot be identified as English or Scottish, a further 2,124 respondents who do not report wage, an additional 505 respondents with occupational fields, which cannot be matched to fields of study, and 371 respondents because of several missing covariates.

²⁷ Note, results from the IEA Third International Mathematics and Science Study (TIMSS) in 1994–95 indicate no significant differences between England and Scotland in the mathematics achievement for students in fourth and eighth grade. There are, however, some differences in the science achievement scores. English students in the eighth grade appears to do somewhat better than their Scottish counterparts, although there is no significant difference for fourth graders.

²⁸ These include: Math/Computer Sciences, Physical Sciences, Architecture, Engineering, Biological Sciences, Health, Social Services, Social Sciences, Business, Law, Education, and Arts.

²⁹ Although a positive sign of the coefficient on *SCOT* would have been more consistent with Proposition 3 (to the extent that an insignificant sign is at all informative).

³⁰ Recall that, with exogenous switching, Proposition 4 predicts higher expected wages loss in Scotland than England when returns to match quality are relatively high. However, with endogenous switching, it is possible for the expected wage loss to be higher in England than Scotland even when the return to match quality is relatively high. And indeed, although insignificant, this estimate is consistent with the simulation in Figure 3.

³¹ Although I do not find that the coefficient on *SWITCH* is affected when controlling for whether individuals changed fields of study *during* university, it may still capture some unobservable traits that differ across individuals.

³² College grades may serve as a useful proxy for these unobservable shocks. However, this is not available in this context because most students in Britain do not take modular courses, each with a separate grade.

³³ Although it is possible that English students wish to choose fields, which facilitate switching in order to avoid specializing in an excessively narrow field, the number of slots in each field in England is essentially determined by government funding.

³⁴ On a related note, there appear to be no significant differences in formal job training between England and Scotland. Nonetheless, there may still be greater informal learning on the job for Scottish individuals to make up for lower levels of skill upon entering the labor market.

³⁵ These regressions are run as ordered probits. Categories include: ‘not at all’, ‘a little’, ‘a lot’, and ‘a great deal’. Similar results are obtained when collapsing these categories into larger groupings.

³⁶ Furthermore, the larger magnitude of average wage losses from switching in England does not necessarily imply lower average wages overall because individuals who do not switch earn higher wages in England than in Scotland.

³⁷ In other words, match quality may be more relevant in determining how much satisfaction students derive from studying and working in a given field. Indeed, this is consistent with Arcidiacono’s (2004) finding that most sorting across majors is due to different preferences rather than differential monetary returns to ability.

³⁸ This can occur when the relative return to match quality at which the probability of switching in the early regime crosses that of the late regime is *below* the relative return to match quality at which the expected wages in the late regime cross those of the early regime.

³⁹ As X_i and Y_i^* already correspond to the mean of the samples, we will use x_i and y_i^* instead of \bar{x}_i and \bar{y}_i^* .

⁴⁰ Note also that the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$ is distributed $N(\mu(\mathbf{x}), w)$ with $w = (\pi + q)\pi q$.

⁴¹ This corresponds to a linear loss function, $L_i(\theta, s) = \theta_{[k]} - \theta$, where $\theta_{[k]} = \max\{\theta_1, \dots, \theta_k\}$ is normalized to zero and with an additional negative cost associated with the amount of sampling from the population i .

⁴² In other words, $\eta^2 = \text{Var}(m(\mathbf{x}))$ where $m(\mathbf{x}) = \int \mu_i(\mathbf{x}|\theta) f(\theta) d\theta$. Wetherill and Ofosu (1974) determine η^2 in an analogous framework — see equation (9.5) on page 263.

⁴³ This equation is equivalent to equation (27) of Dunnet (1960). For the case of $k = 2$, it is also equivalent to an equation derived by Bross (1950) as $E[\max(Z_1, Z_2)] = 1/\sqrt{\pi}$ (see Bose and Gupta, 1959).

⁴⁴ So $Y^{[k]}$ is a random variable representing observations from an extreme value distribution. See Gupta and Miescke (1994, 1996) and Miescke (1999) for similar decompositions.

References

- Afonja B. (1972) 'The Moments of the Maximum of Correlated Normal and T-variates', *Journal of the Royal Statistical Society* 34: 251–262.
- Altonji J. G. (1993) 'The Demand for and Return to Education When Education Outcomes Are Uncertain', *Journal of Labor Economics* 11(1): 48–83.
- Arcidiacono P. (2004) 'Ability Sorting and the Returns to College Major', *Journal of Econometrics* 121(1–2): 343–375.
- Berger M. C. (1988) 'Predicted Future Earnings and Choice of College Major', *Industrial and Labor Relations Review* 41(3): 418–429.
- Bose R. C. and Gupta S. S. (1959) 'Moments of Order Statistics from a Normal Population', *Biometrika* 40: 318–335.
- Bross I. (1950) 'Two Choice Selection', *Journal of the American Statistical Association* 45: 530–540.
- Cohn E. and Kahn S. P. (1995) 'The Wage Effects of Overschooling Revisited', *Labour Economics* 2(1): 67–76.
- DeGroot M. H. (1970) *Optimal Statistical Decisions*, New York: McGraw-Hill Company.
- Department of Employment (1988) *National Survey of 1980 Graduates and Diplomates, 1986–1987* [computer file]. Colchester, Essex: UK Data Archive [distributor], SN: 2314.
- Dolton P. J. and Vignoles A. (2000) 'The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market', *Economics of Education Review* 19(2): 179–198.
- Dolton P. J. and Vignoles A. (2002) 'Is a Broader Curriculum Better?', *Economics of Education Review* 21: 415–429.
- Dunnet C. W. (1960) 'On Selecting the Largest of k Normal Population Means', *Journal of the Royal Statistical Society* 22: 1–40.
- Evans K. (1975) *The Development and Structure of the English Educational System*, London: University of London Press.
- Grogger J. and Eide E. (1995) 'Changes in College Skills and the Rise in the College Wage Premium', *Journal of Human Resources* 30: 280–310.
- Gupta S. S. and Miescke K. J. (1994) 'Bayesian Look Ahead One Stage Sampling Allocations for Selecting of Largest Normal Mean', *Statistical Papers* 35: 169–177.
- Gupta S. S. and Miescke K. J. (1996) 'Bayesian Look Ahead One-stage Sampling Allocations for Selection of the Best Population', *Journal of Statistical Planning and Inference* 54: 229–244.
- Hamermesh D. and Donald S. (2006) 'The Effect of College Curriculum on Earnings: An Affinity Identifier for Non-ignorable Non-response Bias', Mimeo.
- Hunter S. L. (1971) *The Scottish Educational System*, Oxford: Pergamon Press.
- Johnson W. R. (1978) 'A Theory of Job Shopping', *Quarterly Journal of Economics* 92(2): 261–278.

- Jovanovic B. (1979a) 'Job Matching and the Theory of Turnover', *Journal of Political Economy* 87(5): 972–990.
- Jovanovic B. (1979b) 'Firm-specific Capital and Turnover', *Journal of Political Economy* 87(6): 1246–1260.
- McCall B. (1990) 'Occupational Matching: A Test of Sorts', *Journal of Political Economy* 98(1): 45–69.
- McMillen D. P., Seaman P. T. and Singell L. D. (2007) 'A Mismatch Made in Heaven: A Hedonic Analysis of Overeducation and Undereducation', *Southern Economic Journal* 73(4): 901–930.
- Malamud O. (2011) 'Discovering One's Talent: Learning from Academic Specialization', *Industrial and Labor Relations Review* 62(2).
- Miescke K. J. (1999) 'Bayes Sampling Designs for Selection Procedures', invited chapter in Ghosh S. (ed.) *Multivariate, Design, and Sampling*, New York: M. Dekker: 93–117.
- Miller R. (1984) 'Job Matching and Occupational Choice', *Journal of Political Economy* 92(6): 1086–1120.
- Neal D. (1999) 'Complexity of Job Mobility among Young Men', *Journal of Labor Economics* 17(2): 237–261.
- Nelson R. R. and Phelps E. S. (1966) 'Investment in Humans, Technological Diffusion, and Economic Growth', *American Economic Review* 56(1/2): 69–75.
- Osborne G. S. (1967) *Scottish and English Schools: A Comparative Survey of the Past Fifty Years*, Pittsburgh, PA: University of Pittsburgh Press.
- Rumberger R. W. and Thomas S. (1993) 'The Economic Returns to College Major, Quality and Performance: A Multilevel Analysis of Recent Graduates', *Economics of Education Review* 12(1): 1–19.
- Shaw K. L. (1987) 'Occupational Change, Employer Change, and the Transferability of Skills', *Southern Economic Journal* 53: 702–719.
- Sicherman N. (1991) 'Overeducation in the Labor Market', *Journal of Labor Economics* 9(2): 101–123.
- Squires G. (1987) 'The Curriculum' in Becher Tony (ed.) *British Higher Education*, London: Allen & Unwin.
- Trow M. (1999) 'From Mass Higher Education to Universal Access: The American Advantage', *Minerva* 37: 1–26.
- University of Edinburgh (2003) *2003 Undergraduate Prospectus*, Edinburgh: University of Edinburgh.
- Weiss Y. (1971) 'Learning by Doing and Occupational Specialization', *Journal of Economic Theory* 3: 189–198.
- Welch F. (1970) 'Education in Production', *Journal of Political Economy* 78(1): 35–59.
- Wetherill G. B. and Ofosu J. B. (1974) 'Selection of the Best of K Normal Populations', *Applied Statistics* 23(3): 253–277.
- Wiswall M. (2008) 'Licensing and Occupational Sorting in the Market for Teachers', Mimeo.