Abstract

Recent research on economic returns to higher education in the United States suggests that those with the highest wage returns to a college degree are least likely to obtain one. We extend the study of heterogeneous returns to tertiary education across multiple institutional contexts, investigating how the relationship between wage returns and the propensity to complete a degree varies by the level of expansion, differentiation, and cost of higher education. Drawing on panel data and matching techniques, we compare findings from the US with selection into degree completion in Germany and the UK. Contrary to previous studies, we find little evidence for population level heterogeneity in economic returns to higher education.

Introduction

The returns to higher education are consistently found to be large; a university degree is associated with improved life chances across nearly all domains of life. In addition to earning 84% more across their lifetimes than a worker with a high school degree (Carnevale et al. 2011), those with a bachelors degree or higher enjoy better health, have more stable marriages, and report higher life satisfaction (Hout 2012). Recognizing these advantages, individual and state investment in higher education increased dramatically in the past four decades. A recent cross-national study noted that, across the 15 countries under consideration, both eligibility for and attendance in higher education doubled for cohorts born between 1940 and 1970 (Shavit et al. 2007). Despite this increase in the supply of university[1] graduates, the wage returns to a university degree remain sizable and continue to grow (David Card and Lemieux 2001; Long 2010).

However, recent work on the US finds that wage returns to a university degree are heterogeneous; some individuals benefit more than others (J. E. Brand and Xie 2010; Carneiro et al. 2011; D. Card 1999). In order to obtain the greatest returns on the substantial investments of individuals and governments in higher education, it is critical to know not only the average returns to a university degree but also how those returns are distributed across a population. Are those most likely to complete university the ones who stand to gain the most?

Standard economic theories of educational attainment anticipate economically efficient sorting into higher education. Expected wage returns are assumed to be dominant in the decision making process to complete higher education, and according to the principle of comparative advantage, those most likely to complete University will be those with the most to gain net of the cost of completion (Willis and Rosen 1978; Carneiro et al. 2011). Thus, net of costs, we should expect to see a *positive* relationship between the propensity to complete college and wage returns to a degree.

Although this standard theory of positive selection is widely accepted (Hout 2012), some recent research suggests the opposite. In both the United Kingdom and the United States those with *lower propensities* to complete a university degree have been found to gain the most from having one (Dearden et al. 2004; J. E. Brand and Xie 2010). Sociologists have developed a theory of *negative selection* to account for these findings, emphasizing the primacy of non-economic determinants of university attendance and completion that result in the economically inefficient distribution of university degrees.

The question of negative or positive selection is at the crux of theoretical debate over the role of higher education in exacerbating, or ameliorating, intergenerational immobility. If we assume that expanding higher education will increase access to higher education for everyone in accordance with their expected economic returns, regardless of socioeconomic background, then expanding higher education should encourage regression toward the mean in terms of the intergenerational transmission of income and reduce socioeconomic reproduction. If, on the other hand, as expected by sociological theories such as *maximally maintained inequality* (MMI), privileged class members exploit their relative advantage to secure opportunities for degrees independent of their economic returns, expanding higher education may increase socioeconomic reproduction across generations. The question thus has important policy implications: in a world of scarce resources, where state support for higher education is limited to only a certain proportion of the population, it is critical to understand what government policies might create more or less economically efficient sorting into university completion. In the United States, Germany, and the United Kingdom the opportunities to enter higher education, costs of higher education, and alternatives to higher education differ both between countries and within each country across time. In this article we exploit this cross-national and cross-temporal variation to derive and test hypotheses about the institutional mechanisms that encourage and discourage positive and negative selection.

Background

University enrollment and completion is not randomly distributed across any population. There are two possible sources of selection. First, individuals may sort into university completion on the basis of pre-existing characteristics, such as family background, academic ability, and geographic location. For instance, we know that young adults with parents who completed a university degree, who perform well in school, and who live nearby to a university are more likely to attend and complete university (Deil-Amen and Turley 2007).

The second source of heterogeneity receives less attention: heterogeneous *treatment* effects, or the fact that university completion may result in larger wage returns for some individuals and smaller returns for others (J. E. Brand and Xie 2010; Xie et al. 2011). Individuals may sort into university not only based on their background characteristics, but also on the benefit that they will obtain from the degree itself. This is referred to as "treatment effect heterogeneity" (Xie et al. 2011). We ideally want positive selection into university completion. If those who benefit most from completion are the ones who actually complete, then investments in higher education will yield larger returns at a societal level.

Positive and negative selection

The question of whether those who complete university gain more or less than those who do not has sparked a large debate. At the center of this debate is disagreement about the importance of economic cost-benefit analysis in the decision to attend and complete University. Economic theories of positive selection, based on underlying assumptions of revealed preferences and rational choice, assume that absent imperfect information, credit constraints or uncertainty, individuals will attend college in accordance with expected *economic* returns. Both human capital (G. S. Becker 1994) and signaling theory (Spence 1973) assume positive selection at the population level: those most likely to benefit financially from a university education should be more likely to complete, leading to a relatively efficient distribution of education (Heckman et al. 2006; Carneiro et al. 2011; Willis and Rosen 1978). Although there is a large economics literature on the disrupting role of credit constraints and class and racial differences in sensitivity to price differences on sorting into university enrollment and completion (Heller 1997; Hilmer and Hilmer 2012), this research does not challenge the centrality of economic motivation in higher education and is rarely directly applied to the question of population-level selection into university.

In contrast, the sociological perspectives supporting a theory of negative selection places dominance on *noneconomic* calculations in the decision to go to college such that children from some families may decide to complete University even when there is no economic incentive to do so. First, formal educational sorting often begins at a young age, before most students can predict their wage returns. Sorting into higher or lower educational tracks is therefore strongly influenced by non-economic factors, such as class and gender norms (Buchmann et al. 2008), informational inequities (Grodsky and Jones 2007), and social network composition (Perna and Titus 2005). If tracking occurs early, and if participation in primary and secondary level tracks strongly determine University attendance, then educational choices of students are effectively made by their parents rather than their own economic rational choice.

Second, university education can be a pleasurable pursuit, an entrance to a marriage market (Blossfeld and Timm 2003), and a cultural good (Bourdieu and Passeron 1990) as well as an economic investment. In contrast to tertiary level vocational training, university completion in many liberal arts or non-applied science subjects are intended to promote critical thinking or general knowledge but have only indirect links to occupational requirements. Indeed there is some doubt as to whether even these more general capacities are increased through university attendance (Arum and Roksa 2011). As a result, the decision to attend and complete university may be disconnected from anticipated financial returns (although not general returns to overall "utility" in economic theory).

Finally, the perceived risk of university attendance and completion varies across individuals. Some students—regardless of their anticipated benefit—may perceive their exposure to significant debt and lost wages as riskier than their peers. The effect of such information failures can be incorporated into a traditional economic behavioral model. However the non-economic motivation of status maintenance has a role to play in assessing risk as well. Particularly less advantaged students may prefer to maintain their existing class status, forgoing the risk of university attendance and completion. In contrast, students from more privileged backgrounds are more likely to experience downward mobility in status if they choose not to complete university, and hence may be more willing to take larger risks to obtain a degree (Breen and Goldthorpe 1997).

To summarize, whereas most economic models expect economic motivations to dominate the decision to attend and complete university, sociological models anticipate that for many students non-economic motivations will dominate. However, even if we assume that non-economic motivations drive educational decisions, it does not necessarily follow that *negative* economic selection into University attendance and completion should occur at the population level. Rather, as argued by Brand and Xie (2010: 279), non-economic motivations to attend university will need to be highest for those with the lowest economic motivations to create negative economic selection at the population level.

Although the positive and negative selection hypotheses disagree on the overall direction of selection into a college degree, they both anticipate heterogeneous effects. Indeed, recent research from the US and UK demonstrate that returns are heterogeneous (J. E. Brand and Xie 2010; Carneiro et al. 2011; Dearden et al. 2004). Brand and Xie (2010) utilize propensity score matching to simultaneously account for multiple sources of heterogeneity in the income returns to academic post-secondary education. They argue, in the US case, that the propensity to complete university and income returns are negatively associated. Using similar methods with the British Cohort Study, Dearden et al (2004) report a nonlinear relationship, finding that individuals in the middle of the propensity score distribution reap the largest rewards. By contrast, Carneiro et al (2011) combine matching techniques within a latent class model of the desire to attend college in the United States and show that students positively select into university.

Relying on different data sources in different countries, different methods, and different age cohorts, results from the studies above are intriguing but essentially incomparable and hence the debate over the direction of selectivity continues (Hout 2012). Moreover these single country analyses implicitly assume that the selection process unfolds in a similar pattern independent of time and space.

Such investigations leave untouched the institutional mechanisms behind both positive and negative selection. Yet the decision to attend and complete university is dependent on the educational policies that vary across countries and across time. Not only do education policies such as tuition fees and length of study directly alter the cost-benefit calculus of university completion, they can also have indirect impacts on the importance of non-economic factors discussed above. Decreases in tuition, for instance, might reduce the perceived risk of higher education among less advantaged students, and subsequently reduce inequalities surrounding class norms in higher education participation. Similarly, expanding the number of higher education institutions may reduce competition for university slots and hence weaken the link between access to University and previous track placement. In the section that follows, we draw on recent comparative research on the link between higher education and social stratification (Shavit et al. 2007; Van de Werfhorst and Mijs 2010; Malamud and Pop-Eleches 2011) to develop hypotheses on the relationship between three institutional differences and positive or negative selection into university at the population level.

Institutional Context

Shavit et al. (2007) identify three characteristics of higher educational systems which shape the relationship between higher education and social stratification: higher education eligibility and attendance, the mode of differentiation, and the market structure. Higher education eligibility and attendance broadly capture the level of higher education expansion. Eligibility is the percentage of a cohort who obtains the necessary secondary qualification that allows access to university education; attendance is the percentage of the cohort who actually attends university. Modes of differentiation refer to the kinds of higher education offered: unified, binary, and diverse. Unified higher education systems primarily offer university education that is theoretically based and allows access to skilled professions. Binary systems offer university education alongside a secondary tier of tertiary level vocational training, which is more applied yet may still be prestigious and of high quality. Diversified systems offer a secondary tier which provides a mix of vocational and academically orientated courses. Finally, the market structure is the degree to which the costs of higher education are borne by the state versus borne by students and their families. Across many developed countries, expansion tends to be positively associated with diversification and marketization of higher education.

Drawing on the theories of *maximally maintained inequality* (MMI) and *effectively maintained inequality* (EMI), Shavit and coauthors explore how different constellations of these three factors are associated with the reproduction of educational inequality across generations. MMI posits that even as higher education systems expand to accommodate a larger percentage of a cohort, advantaged population members will maintain their relative advantage until the point of saturation: in other words, until essentially all advantaged families secure a place in higher education for their children, expanding higher education will do little to decrease inequality (Adrian E Raftery and Michael Hout 1993). Similarly, the EMI hypothesis is that the increased diversification that often accompanies expansion will lead to diversion effects, such that less advantaged population members are diverted from more prestigious educational offerings in favor of lower prestige secondary tiers (Lucas 2001; Hillmert and Jacob 2003). Cross-national and cross-temporal tests offer support for both these hypotheses, generally finding that inequality in higher education is maintained even under expansion but not necessarily to the degree hypothesized under MMI, and that the relationship between expansion and inequality differs across contexts (Paterson and Iannelli 2007; A.E. Raftery and M. Hout 1993; Shavit et al. 2007; Tam and Jiang 2014; Bar Haim and Shavit 2013; R. Becker and Hecken 2009).

In this article, we borrow these typologies to develop hypotheses about the relationship between institutional structure and economic selection into university, a related but distinct research area.

Eligibility and Attendance

Higher education eligibility and attendance are influenced first by the primary and secondary school tracking systems and second by the supply of higher education available. Educational choices made in systems that track students earlier in the life course are more likely to be influenced by class background (Hanushek and Wößmann 2006). We would also expect that in systems where tracking is fluid and access to university is allowed through multiple channels, educational choices will be more strongly linked to a student's own ability and economic choices than in systems where tracking is rigid, with little cross-over opportunities between tracks. As students age they (and their parents) receive greater information about their own abilities and future plans, allowing for more informed choices about higher education; flexible tracking systems will enable students to capitalize on improving information and make more economically efficient choices.

Higher education eligibility and attendance will also be influenced by the degree of educational expansion and total number of slots available – both in upper level secondary tracks as well as in higher education institutions themselves. On the one hand, limited higher track and university slots may positively influence selection. Positive selection theory would expect that limiting higher education eligibility and attendance to strong performers will result in a student body most able to translate their human capital into higher wage returns. On the other hand, if pre-university performance is largely an indicator of socioeconomic background, strongly “meritocratic” selection mechanisms may only serve to reproduce class inequalities, leading to greater negative selection into university completion and diminished labor market returns. In this case, more expansive and accordingly less competitive higher education systems may increase positive selection.

Mode of Differentiation

Higher education can encompass both academic and vocationally orientated tracks. Sociologists argue that, in the United States, university completion has become a norm for youth from more advantaged backgrounds, resulting in weaker economic selection amongst those with the highest propensity to complete (J. E. Brand and Xie 2010; Beattie 2002). As argued above, the relatively weak links between academically orientated higher education and many occupations may also weaken the importance of economic motivations relative to non-economic motivations in the decision to attend and complete university. In contrast, vocational training is directly linked to work and wage returns (Ainsworth and Roscigno 2005), and as it is generally less prestigious, may convey fewer non-economic benefits. Vocational programs are also usually shorter than academic degrees, and frequently combine work experience with training, both of which lower costs and foregone wages. Hence, we would expect that higher education systems with developed vocational training – binary systems – would have more positive selection than unitary systems. On the other hand, binary systems and still more diversified systems have been shown to “divert” disadvantaged youth away from academic tracks (Lucas 2001; Hillmert and Jacob 2003). If these youth would have enjoyed greater economic benefits from university rather than vocational programs, binary systems may increase negative selection specifically into academic orientated higher education, even as higher education in general may become more positive selected.

Market Structure

Both the positive and the negative selection hypotheses predict that credit constraints are likely to make sorting into higher education less economically efficient. Lower costs will thus increase positive selection by allowing those who stand to benefit most to attend, regardless of financial position. Moreover, changes in tuition may also influence inequality in the non-economic motivations to attend. As emphasized in the negative selection hypothesis, if differential access to information on the costs and benefits of higher education cause the creation of norms of university attendance among some groups but not others, lowering costs should equalize perceived costs and benefits and thus promote a more equal culture of education among disadvantaged and advantaged groups alike.

In the empirical analysis to follow, we build on the groundbreaking work in the US context by Brand and Xie (2010) to compare the heterogeneous wage returns of a university degree across two additional country contexts, Germany and the United Kingdom. To better isolate the effect of institutional differences between the countries, we restrict our analysis to men born between 1955 and 1965, observing their earnings in 2000 and 2004. These two countries, for these birth cohorts, were chosen both for comparability of data across countries and for their variation in the characteristics outlined above. Because family, work, and education interact in important ways and are often endogenously determined for women, we also chose to simplify the analysis by focusing exclusively on men (Kalleberg 2000). Examinations of cross-national differences across time, and across genders, each merit their own papers and space constraints prohibit us from pursuing them here.

We briefly summarize each country context for the relevant age cohorts below and describe how the unique characteristics of each context relate to the hypotheses drawn from the positive and negative selection theories.

Country Contexts

Early rigid tracking and a highly developed vocational track characterize the German educational system. Between ages 10 and 12 teachers, parents, and students together decide which of three educational tracks students should enter. Only the highest of the three levels directly prepares students for entrance into university. Students attending the lower tracks typically continue in vocational education and/or an apprenticeship. In 2006, approximately 40% of German secondary students were enrolled in an academic program, while the other 60% were enrolled in vocational programs. Transition to academic tertiary education is highly selective – in the same year, 35% of German secondary students transitioned into tertiary education, compared to 57% in the UK and 64% in the US (OECD 2011). Most universities in Germany charge nominal fees and are state run. Additional cost of living assistance is available to all tertiary level students; prior to 1983 cost of living for lower income families was state supported and after 1983 half of this support was to be repaid after graduation. The number of years it takes to attain a bachelor's degree depends on the subject studied; however, the average time to degree is 6 years, higher than in the US or the UK.

The United Kingdom (excluding Scotland and Northern Ireland) is characterized by weaker tracking and fewer opportunities for vocational training. Compulsory education ends at 16; students may choose to continue with their academic education and complete A-Level exams (required for university entrance), enter a further education institution to pursue vocational training, or leave school completely. Almost 60% of upper secondary students continue with the academic program, and only 31% of UK youth enroll in vocational upper secondary education. The 1962 Education Act guaranteed tuition and maintenance support for all students pursuing higher education. This support declined in real value throughout the 1980s, and in 1990/1991 a student loan scheme was introduced and maintenance grants phased out (Cheung and Egerton 2007). University attendance and completion increased especially rapidly in the UK in the past decade; in 2009 45% of 25-29 year-olds in the UK completed a university degree, as compared to only 26% in 2000. A bachelor's degree typically takes 3 years to complete, the shortest course of the three countries.

In contrast to Germany and the UK, which are binary systems offering two main types of post-secondary education, vocational and academic, the US system is characterized as *diversified* (Meek et al. 1996). The US includes in its lower post-secondary strata programs that are vocational as well as programs that lead to academic qualifications such as two-year colleges. Due in part to the plethora of options, the US university system affords greater access, with high enrollment rates. However, the relatively high tuition costs and lack of public support for maintenance costs lead to lower conditional completion rates (Lau 2003) than in the UK and Germany. Although the US was a leader in university attendance and completion rates throughout the past four decades, expansion has flat lined in the past ten years, especially in terms of completion, and as of 2009 40% of 25-29 year olds completed university. Vocational training in the United States is also weaker and less often linked to apprenticeship and job opportunities than in Germany (Shavit and Müller 1998).

This brief overview of the three country systems suggests three major areas of differentiation between the country contexts. These differences are summarized in table 1. First, the British and German systems are binary, and the US system is diversified. Germany has by far the most developed vocational system, and in the US this system is essentially absent. The UK represents something of a middle ground with weaker vocational training than Germany but a stronger system than in the US, particularly for the cohort studied here.

The second area of differentiation is the cost of a university degree. As table 1 shows, the cost of a university degree varies widely across the three countries. For instance, in 1976 the average cost of tuition, and room and board at a public university in the US was $6,877 (in 2006 dollars); by contrast, higher education was essentially free for students in the UK and Germany during this same time period. However, the lost wages due to a longer average time to degree sets Germany apart from the UK. In terms of costs, then, the US stands out as the highest cost country, followed by Germany, and finally the UK, with the lowest opportunity and direct financial costs for higher education.

The third area of differentiation is university eligibility and enrollment. At only 16% of the 45-54 cohort with a university degree, Germany stands out again as the least expanded university system; however the United Kingdom was also highly selective for the cohorts under observation here. Eligibility also differs strongly across the countries, whereas 89% of US men ages 45-54 obtained the high school degree necessary for university access, less than one in four British or German men were eligible for university at that age. Relatively low numbers of students take the exams —A-levels in the UK or Abitur in Germany—necessary to attend university, self-selecting prior to the entrance exams rather than in response to them. In contrast, given the diversity of universities and colleges with different entrance requirements, selectivity is less tightly linked to former academic performance in the US. Finally, linking eligibility, enrollment and completion, it is important to note the different university retention rates across the three countries. Whereas 75 and 79 per cent of those who start a university program receive a degree in Germany and the UK, respectively, only 64 per cent of US college students receive a degree (Quinn 2013).

[TABLE 1 ABOUT HERE]

Drawing from the review of the negative and positive selection hypotheses outlined above, as well as the sources of variation between the countries and cohorts under observation here, we set forth the following competing hypotheses.

The first set of hypotheses address a general trend:

H1A: Economic rational choice models expect economic motivations to dominate the decision to attend and complete university; hence we should observe positive economic selection into university completion.

H1B: Sociological models expect that non-economic motivations may dominate the decision to attend and complete university, and that non-economic motivations will be strongest for those with the weakest economic incentive to complete university. Hence, we should observe negative economic selection into university completion.

The second two hypotheses address differences between countries:

H2A: Economic driven decisions for university completion should be *even more positively selective* in less expansive educational systems. Hence, we should expect more positive selection in Germany or in earlier cohorts in the UK.

H2B: More expansive education systems confer less prestige and status. A system which affords greater access for everyone should therefore be less driven by non-economic motivations and university completion should be more positively economically selected. Hence, the USA and later cohorts in the UK should have more positive selection.

H3A: Binary systems offer more options for investment in human capital; binary systems should therefore create more efficient economic sorting into higher education. Hence, Germany and earlier UK cohorts should have the most positive selection.

H3B: Binary systems will divert less socioeconomically advantaged students from prestigious tracks, resulting in the greater influence of non-economic motivations in university (versus vocational) post-secondary completions and less efficient economic sorting. Hence, Germany and earlier cohorts in the UK should have less positive selection.

H4: Lower costs should reduce market distortions due to credit constraints and informational inequalities, resulting in more efficient economic sorting. The early cohorts in the UK and Germany should therefore have the most positive selection.

In balance, economic theories of selection into university completion expect more positive selection in early cohorts in the UK and Germany, whereas non-economic theories of selection into university completion expect less positive selection in early UK cohorts and Germany.

Methods

Homogeneous effects

We begin with a simple analysis of the wage returns to a university education in each country separately. The level of an individual's wages (*W*) is a function of whether that individual completed university (*D*), a series of covariate controls such as their sex, socioeconomic background, and age (*X*), and an error term (*u*), or:

$$W_i = \beta \cdot D_i + \gamma \cdot X_i + u_i \qquad (1)$$

where β represents the estimated average effect of completing university on wages. We make two important assumptions about this model in order to estimate the unknown parameter β by an OLS regression of wages on completed university and controls, where the treatment (completing university) is not randomly allocated but rather is self-selected. First, we assume that those who attend and complete university are not different in unmeasured ways from those who do not. Second, we assume that those who complete university, *D*=1, all receive the *same* benefit for completion.

Propensity score matching for heterogeneous treatment effects

Next, we relax the second assumption to ask whether and how the effect of university attendance on wages may be heterogeneous across different populations. To do so, we employ the propensity score approach used by Brand and Xie for the United States, which first creates balanced strata within a population based on their propensity to complete university and then relaxes the assumption of homogeneous returns across the propensity strata. The central goal of this paper is to compare Brand and Xie's results from the USA with similar analyses in Germany and the United Kingdom to examine institutional differences in the direction of selection into university completion. In so doing, we also provide a replication of the original Brand and Xie manuscript while assessing the application of the method in new contexts.

Following Xie, Brand, and Jann (2011) we first estimate individuals' different probabilities, *p*, of completing university, *D*, given a similar vector of demographic and socioeconomic covariates *X* as above.

$$P = p(D_i = 1|Z) \quad (2)$$

If we assume that the vector of covariates *X* captures all of the ways that those who complete university are different from those who do not complete university, then we can assign each individual *i* a propensity score *P* that describes his likelihood of completing university conditional on *X*.

After the assignment of propensity scores, individuals are grouped based on propensity scores into strata (i.e. a group of individuals with similar propensities to complete university). Within a propensity score stratum, some individuals completed university while others did not. Within each stratum, no significant differences (at the .01 level) in the average values of covariates (*X*) between university completers and non-completers are allowed. This condition, referred to as balance, ensures comparability between completers and non-completers on measured covariates. Given a rich set of covariates, it is hoped that the stratification by the propensity score is an effective way to remove most biases between the treated and untreated groups.

Using these propensity score strata we estimate a hierarchical linear model predicting wages. The first level comprises a series of regressions of wages, *W,* on college completion *D* for each strata separately. Doing so allows the effect of college completion for observably identical individuals to differ across each strata *j*:

$$W_{ij} = \beta \cdot D_{ij} + u_{ij} \quad (3)$$

Next, we evaluate a trend across the strata using variance-weighted least squares regression of the strata-specific treatment effects β, obtained in equation (3), on strata rank at level-2. This step departs from the conventional use of propensity scores in constructing strata, where the emphasis is usually on removing biases due to covariate imbalances simply by averaging the estimated treatment effects across strata (Xie et al. 2012). Instead, the main research objective we emphasize is to look for a systematic pattern of heterogeneous treatment effects across strata. Following Brand and Xie (2010), we model the heterogeneity pattern as a linear function across strata ranks *j*, using the standard errors for β obtained in equation (3) to weight according to variance of the estimates:

$$\beta = \gamma \cdot j + u$$

A negative trend γ suggests *negative selection* into university: individuals who are least likely to complete university, but do complete, receive greater economic returns compared to individuals within their stratum who did not complete, and this difference in economic returns is larger than the difference in economic returns among individuals who are more likely to complete university. A positive trend suggests *positive selection:* among those most likely to complete university, the difference in wage returns between completers and non-completes is larger than the difference among those least likely to complete. No significant effect at level 2 suggests that the returns to a university degree do not vary depending on propensities to complete university.

Although we are concerned with heterogeneity in the effects of university completion on wages, this propensity stratification and multi-level approach to analyzing propensity scores imposes homogeneity *within* propensity score strata. In other words, we assume that, within a propensity score stratum, all respondents will experience the same benefit (or non-benefit) to university completion. We understand that this is a stringent assumption. Propensity score models cannot account for unobservable characteristics, and are therefore still susceptible to omitted variable bias. However, by utilizing very rich, longitudinal data, including measures of cognitive and academic ability normally unobserved, and allowing for nonlinearity between observables and the outcome, we can more closely simulate an experimental setting with control and treatment groups than is usually the case. Moreover, by examining men only, close in age, primarily native and non-minority, we hope that we get closer to this homogeneity assumption (Henderson et al. 2011) but we recognize that any effect of education on wages within these strata is the result of both the university degree *and* any unmeasured characteristics.

To correct for missing data, we use the MI suite in Stata 14 to multiply impute 20 datasets with complete information. Propensity score matching models are then fitted using user written commands pscore (S. O. Becker and Ichino 2002) and pbalchk (Lunt 2013) on each imputed dataset, and the results are combined before estimating wage differences within strata adjusted for multiply imputed data. Finally, we use variance weighted least squares to estimate the level 2 slope adjusting for error in the level 1 regressions.

Limitations of propensity score matching

The propensity score matching approach presents a possibility for causal tests in the absence of experimental data or valid instruments which can be used to approximate random assignment. However, the method rests on other assumptions which may be more or less tenable depending on the research question and data used. A common critique of this method is that the ignorability assumption is likely to be violated. Those who complete university are fundamentally different on unmeasured characteristics compared to those who do not complete precisely because they made it through the college attendance and completion process successfully. Ignoring the influence of these unobserved characteristics will lead to biased estimates (Zhou and Xie 2016).

The severity of omitted variable bias will vary depending on the process of university attainment in a particular country as well as the available data. Propensity score models work best when large numbers of treated and untreated individuals can be matched across a very rich array of characteristics. The first factor in the plausibility of the ignorability assumption is therefore the richness of the available data on which to match respondents: the more variation that can be captured in the propensity score, the lower the probability of omitted variable bias. The second factor relates to the ability to identify enough "matched" individuals for comparison: small datasets, or very rare treatments, will make matching difficult across the entire propensity range. On the one hand, large datasets can help overcome shortcomings in the method associated with rare treatments: for instance, even if an event is relatively uncommon (such as university completion in the UK for the cohorts studied here), a large enough sample can still ensure a sufficient number of treated and untreated cases with similar propensity scores. On the other hand, where the sample size is smaller, if the event is more common (for instance, university completion in the USA), it may still be possible to match enough individuals because there will be a more even number of treated and untreated cases. In the results to follow, we will show that the method is more defensible in the UK or US contexts than with our German data, and will discuss the meaning of these methodological issues in our discussion.

Data

To maximize comparability across countries, we rely on several data sources with substantial overlap in design and time period with the original Brand and Xie study. These data are described below.

National Child Development Study (NCDS)

In the United Kingdom, we use data from the NCDS, a longitudinal study that surveyed all children born in one week in March of 1958 in England, Scotland and Wales (N=17,634). Eight waves of data have since been collected with more than two thirds of original participants retained in the most recent wave (2010). Given the focus on socioeconomic background, social development, and educational outcomes at younger ages, and school leaving, the labor market, marriage, and fertility at older ages, as well as cognitive ability tests and detailed achievement information, the NCDS provides ideal data to study the income returns to a university degree. To invoke time-order reasoning for causality, we rely on variables observed in adolescence and early childhood to predict university completion, and focus on wages when respondents are at ages 42 and 46 in 2000 and 2004 (waves 7 and 8).

National Longitudinal Study of Youth

To replicate Brand and Xie (2010) we use the same data, the 1979 National Longitudinal Study of Youth (NLSY). The NLSY includes a nationally representative sample of individuals born between 1957-1965 (N=12,686). The sample of respondents was interviewed annually through 1994 and then every other year since. In 2010-11, 7,565 individuals responded to the survey, for a retention rate of 76 percent. With a focus on important life-course transitions, these data provide an abundance of information on respondents throughout their life course, including cognitive ability tests, academic achievement, and socioeconomic background in adolescence, and educational attainment, occupation, and family in adulthood. We replicate all of the NLSY findings for Brand and Xie (see Appendix 1) but to maximize comparability with the other surveys, focus on wages when respondents are 35-38 in 2000 and 39-42 in 2004.

German Socio-Economic Panel (GSOEP)

The GSOEP is a representative yearly panel survey of 11,000 households, beginning in 1984 in Germany. For comparability with the NLSY and the NCDS, we restrict our analysis to men born between 1955 and 1965, observing their wages in 2000 and 2004. We use the original sample representative of West Germans, the refreshment samples E (1998) and F (2000), and the high income sample 2002. Most of this birth cohort entered the sample as adults and hence observations in adolescence are not available, we therefore rely on retrospective accounts of parental status, school performance, and socialization in childhood from the BIOSOC module implemented in 2000 (Goebel 2015). It is well known that retrospective data can be less reliable than contemporaneous data sources, however educational information is generally found to be one of the most reliably recalled characteristics (Dex 1995). Attitudinal information is less reliably recalled, but we prefer these measures to contemporaneous indicators given that such reports may be endogenous to our treatment variable of interest, university attainment.

Samples

Following Brand and Xie, we focus on the sample of individuals who 1) are at risk for completing a university degree, and 2) are employed but not self-employed at the time we observe their wages. In the first case, we exclude those who are highly unlikely to obtain the qualifications necessary for entrance into a university: high school drop-outs in the US, those with no credentials (GCSEs/O-levels) in the UK, and those who attended the lowest secondary track in Germany. This decision is practically necessary, because so few individuals with very low levels of qualifications obtain a University degree, prohibiting propensity matching. However, we acknowledge that it creates some selection in our sample by omitting the most disadvantaged who may stand to gain the most from a university degree. Given the large number of comparisons conducted both within and across countries, we also chose to simplify the analysis by focusing exclusively on men.

In the case of the UK, we further exclude Scotland because its school system differs significantly from the rest of the country. In Germany, we similarly limit the sample to men born in West Germany. For the birth cohorts under consideration here, both West Germany and the UK had very few native born minorities or individuals with foreign born parents. We therefore omit non-white British from our sample in the NCDS (N=14) and those with a foreign born parent in the GSOEP (N=39). In our replication of Brand and Xie, we follow the sample limitations as outlined in their paper (2010): we limit the sample to those who completed high school by 1990, took the cognitive ability test before completing high school, and are not missing information on any of the variables used in the analysis. The resulting sample size is 838 in 2000 and 770 in 2004. In contrast, for the German and UK samples, we use multiple imputation to create complete samples of (non-self) employed men with non-zero wages. The final sample size for the UK is 2682 in 2000 and 1911 in 2004. For Germany the final sample size is 700 for 2000 and 685 for 2004. Because of the inclusion of refreshment and high income samples in 2000 and 2002, the German sample in 2004 only contains 548 cases that were also present in 2000.

These restricted samples are more homogeneous in their returns to a degree than the total populations of the countries under consideration (Henderson et al. 2011). Extensive sensitivity tests using alternative samples are conducted for the German and the UK analyses, including the inclusion of those with 0 wages and the self-employed, and for the German case we conduct a full replication with an alternative data set, the National Educational Panel Study (NEPS). These are described at greater length below and the results reported in Appendices A2-A3.

Measures

Table 2 provides a description of the measures used to predict the propensity score strata for completing university in the NCDS and the GSOEP. We also include for comparison the variables from our replication of Brand and Xie with the NLSY. This set of variables including parental education and income, academic achievement, cognitive ability, parental encouragement, and tracking, is widely used to predict both educational attainment and occupational status (A.E. Raftery and M. Hout 1993; McLanahan and Sandefur 1994; Hauser et al. 1983). Although the measurement of most variables is straightforward, we discuss a few harmonizing issues across the surveys, and variations in measurement.

[INSERT TABLE 2 ABOUT HERE]

Several variables are used to capture respondents' social background including parental education, income, and occupation. Weekly family income when the respondent was age 16 is included in the NCDS analyses, but unfortunately not available in the GSOEP. Because occupational status plays a vital role in socioeconomic status, the development of norms, and access to educational opportunities, we include a series of dummy variables to indicate whether the respondent's father was in a managerial or professional occupation, a skilled occupation, or an unskilled occupation in the NCDS, and we use international socioeconomic index (ISEI (Ganzeboom and Treiman 1996) scores (ranging from 16 to 90) for father's occupation in the GSOEP. An indicator of whether the respondent lived with two parents in childhood is also included in both surveys.

Respondents' scores on standardized math, reading and generalized ability tests, capture high school academic achievement in the NCDS. In the GSOEP, respondents' self-reported grades at the time they left high school are used to measure academic achievement. To control for tracking and eligibility, we include in the GSOEP an indicator for whether the respondent passed the Abitur test, required for entrance into an academic university program. In the NCDS, those who have completed at least 2 A-level exams are considered to be eligible for University.

Brand and Xie (2010) emphasized the importance of controlling for cognitive ability. They used the Armed Services Vocational Aptitude Battery (ASVAB), a battery of 10 intelligence tests measuring knowledge and skill in areas such as mathematics and language. This test was administered to all respondents prior to high school completion, and is a widely used intelligence test used throughout the US (Hunter et al. 1985). The NCDS uses a general ability test to measure intelligence. It includes a verbal and non-verbal component. In contrast to the ASVAB, the general ability test was developed for the survey when respondents were 11 years old, however it correlates highly with IQ-tests administered in late adolescence (Douglas 1964). Unfortunately, there is no similar cognitive test prior to University completion in the

GSOEP. This, along with the smaller sample size and rarer university completion in Germany, presents particular challenges for using propensity score matching for this context.

Dependent Variable

In the level 1 HLM analysis, we predict respondents' wages. Wages are measured as gross monthly wages and standardized to 2000 Euros. Following convention we take the logarithmic form for analysis. We focus on mid-career wage returns observed in 2000 and 2004, when respondents are between the ages of 35-46 in all samples.

Independent Variable

In the NLSY, following Brand and Xie, university completion is restricted to those with a degree by 1993, when respondents were 28-31. In the NCDS, similarly, we restrict university completion to those with a completed degree by 1991, when respondents were 35 years old. In the GSOEP, year of degree completion is not reported. Thus, a university completer is any respondent who reports a university degree during the year of observation (2000 or 2004).

Results

Homogenous Effects

Table 3 presents the results of a standard regression analysis predicting the wage returns to a university education in the United Kingdom, Germany, and the United States. Consistent with past research, we find positive (although varying) wage returns to completing a university degree. In the UK, those with a university degree make 6% ($e^{.063}$) more in gross monthly wages in 2000 and 17% more in 2004 compared to observationally identical men who did not complete a degree. These benefits are similar in the US and Germany: our replication of Brand and Xie with 2000 and 2004 NLSY data finds a man with a university degree in the US makes 28% more in 2000 and 17% more in 2004; in Germany, the difference is 15% in 2000 and 35% in 2004.

[INSERT TABLE 3 ABOUT HERE]

Heterogeneous Treatment Effects

These estimates assume that all men experience the same benefit from a university degree. To relax this assumption of homogeneous treatment effects, we carry out the propensity score analysis described above, matching those who complete university with those who do not complete university on key background, achievement, and behavioral characteristics. Following Brand and Xie in the US case, this matching resulted in five balanced strata. We show the results of extending this procedure to the UK and Germany in table 4. In the UK, eight strata were required to achieve balance. In the GSOEP, matching in 2000 resulted in four strata (the two bottom strata were combined to ensure at least 9 treated) and in 2004 in five strata. Balance was established first using pscore, which uses a two step algorithm to create intervals from propensity scores by ensuring first (a) that treated and control groups do not differ in their average propensity score within strata and that (b) that the means of each measured characteristic (Z) do not differ significantly between treated and untreated within each strata (Becker and Ichino 2002: 360). To ensure balance across all imputed datasets, Stata command pbalchk was also used to visualize the matching and to test for balance using the standardized difference between characteristics as well as the multivariate distance. The graphics resulting from this visualization are shown in Figures 1-3, which display the covariate differences between treated and untreated respondents before and after creating the propensity strata. Those in the first stratum have the lowest propensity to complete university while those in the highest strata have the highest propensity to complete university.

[INSERT FIGURES 1 -3 HERE]

Table 4 shows how these strata line up with the observed characteristics of respondents in each of the three datasets. We separate means by whether respondents graduated university or not. In general, the matching process is successful. First, the differences between those who do not complete university and those who do complete university, in each stratum, are minimal. Across each survey within each stratum, balancing criteria are met. Second, as we move across strata, socioeconomic advantage and achievement increase. For example, among the non-university completers in stratum 1 in the UK, 34% of fathers work in unskilled occupations and have on average less than 10 years of education; in contrast, among non-university completers in stratum 8 —those most likely to complete university—only 14% of their fathers work in unskilled occupations, and they have 13 years of education. This descriptive link between parental background and propensity to complete university is similar in Germany. University non-completers in the lowest propensity strata have fathers with 11 years of education and occupational status scores in the upper 30s (production workers and laborers) whereas non-completers in the highest strata had scores on average in the high 50s (lower level professionals, technicians) and 14 years of education.

[INSERT TABLE 4 ABOUT HERE]

Despite the fact that the matching process is successful, the size strata and the number of the treated and untreated cases within each strata differs strongly across the samples. In the larger UK dataset, the strata all include at least 100 cases, even though this data has the richest array of covariates and the resulting highest number of strata. In Germany, in contrast, the middle strata in particular are smaller, with less than 50 total cases, and in the lowest and highest strata there are only a small number of cases with, and without, university completion respectively. This is due to fact that the GSOEP, as a representative household panel study rather than a cohort study, only had a smaller number of cases in the specific age range under study here; the problem of smaller number is further compounded by the fact that university completion is also a rarer event in Germany than in the UK or USA at this time. For this reason, we are more cautious about interpreting the results from the German sample.

In table 5, we turn to our central research question, asking: are those who are most likely to complete university those who enjoy the greatest wage benefits? A negative slope suggests *negative* selection; those who gain the most from a university degree are the least likely to graduate, whereas a positive slope suggests positive selection.

[INSERT TABLE 5 ABOUT HERE]

Starting with the US results from the NLSY, we attempt to replicate Brand and Xie (2010). As can be seen in table 5 as well as our full replication results in Appendix A1, we find statistically insignificant level 2 slopes in both 2000 and 2004 (and for all NLSY survey years). This is in line with Brand and Xie, who also report statistically insignificant level 2 slopes for all survey years they report. However, in contrast to their results, the direction of level 2 slopes we estimate are positive (with one exception in 1994, see Appendix 1). However, the confidence intervals in both Brand and Xie's analysis and our own estimates are large, and the slope coefficients small, and so neither analysis can statistically confirm population level selection for mid-career men in the USA.

Among mid-career men in the UK in 2000 and 2004, we find negative level 2 slopes. In both survey years, the wages of non-completers in the highest propensity strata are higher than those who do not complete. In 2004 (age 46), the wage returns to a university degree are highest in strata 2 and strata 3 and statistically significant among those with the lowest propensities to complete (propensities under 0.4). In contrast among higher propensity groups there is no significant wage difference between university completers and non-completers. For 2004, the negative level 2 slope is also statistically significant at the p<.05 level. Of course, because the number of observations identifying these effects is small in some cases, the standard errors are larger than would be expected from a standard regression approach. Moreover, the statistically significant slope in 2004 is strongly driven by the *negative* association between a degree and wages among those in the highest stratum in 2004, a counterintuitive finding. Still, for the UK, this analysis provides (weak) evidence of negative selectivity for a university degree. The bottom of table 5 summarizes the linear trend across strata, and we see that for the UK, moving from one stratum to the next highest *decreases* the returns in expected hourly wages by 4%. Thus, although essentially all UK university graduates benefit from their degree, those least likely to complete stand to benefit the most.

Turning to the results for Germany in the third column, we see mixed results across the two survey years, and statistically insignificant level 2 slopes in both survey years. In 2000, individuals in the first stratum who completed university earn 27% more than those with a similar (low) propensity to complete university who did not complete. By contrast, those with the highest propensity to complete university (in stratum 4) who in fact completed earn only 2% more than those with a similar propensity who did not complete. Although the general trend in 2000 is negative, it is not significant at the .05 level. This trend furthermore is not observed in 2004. In this survey year, both the first and the fifth strata show the largest gains to a university degree, whereas the middle strata gain less and the wage difference between completers and non-completers in not statistically significant. The level 2 slope for 2004 is statistically insignificant and positive. Similarly to the US, there is no evidence of population level selection among mid-career men in Germany.

These results for the UK and Germany are further illustrated in Figure 4. The negative trend across propensity score strata in the UK is apparent in both survey years. As expected from the results in table 4, the German results are less consistent and less precisely estimated. The middle propensity strata 3 appears to gain the most in 2000, whereas the upper and lower strata gain the most in 2004; in both survey years, there is no discernible downward or upward trends.

[INSERT FIGURE 4 ABOUT HERE]

These results are surprising in two important ways. First, following Brand and Xie, we anticipated observing negative selection in the US case. However, closer inspection of their own findings and our replication exercise showed no statistically significant evidence of selection in either their analyses or our own. Given that all of the level 2 slopes we observe in both are statistically insignificant from 0, the small differences in size and direction of slopes observed between their initial analyses and our own replication attempts do not suggest any substantive difference in observed population level selection. Thus, even for the focal country of the Brand and Xie analysis, and the majority of the literature on selection into university, there is no clear evidence pointing to either the confirmation of positive economic selection (H1A) or negative economic selection (H1B).

Second, following our review of institutional differences between Germany, the UK and the USA, we also expected that selection in Germany and the UK would either be more positive (H2A and H3A), or more negative (H2B and H3B), than selection in the USA. The hypothesis of positive economic selection into university completion would anticipate that in the UK and Germany, both binary systems with more restricted access to university, university completion would be still more positively selected than in the USA. In contrast, the hypothesis of negative economic selection anticipates more positive selection in the USA due to its more expansive policies and variegated options, which should decouple university completion from elite status and prestige. Yet we found neither to be the case: only in the UK was a statistically significant slope documented.

Unfortunately, it is always difficult to explain and defend null findings. As we have described above, particularly the German data was lacking a key characteristic – cognitive ability in adolescence – and suffers from the smallest sample size. To reassure ourselves that our results are robust to a variety of different specifications, we therefore conducted extensive sensitivity tests. In Appendix A1, we show the results for our replication of Brand and Xie with NLSY data for all survey years, demonstrating that our different results are not the choice of the survey years displayed in table 5. In Appendix A2, we show the German results using both the NEPS 2010 (the first wave of data available) as well as the GSOEP 2012 for comparability. We show similarly slightly negative level 2 slopes that are not statistically significant in both datasets for these survey years. In Appendix A3, we show sensitivity results for both the German and UK surveys for two alternative specifications: including controls for marriage and children from the level 2 regressions, and including those with 0 wages and the self-employed. As can be seen in all respective appendices, the null population level finding (for Germany) and the small, but statistically significant negative population level slope (for the UK) are found for all specifications.

Auxiliary Analysis: Educational Expansion

Whereas the US has the most expansive and least selective academic post-secondary university system, university attendance for the same birth cohorts in the UK and Germany was fairly uncommon. Hypotheses H2A and H2B anticipated that differences in expansion would influence selection between these countries yet we found inconsistent support – more negative selection in the UK but no difference between Germany and the US. A better way to isolate the effect of expansion on selection, however, is to evaluate the same country across time, which allows us to control for other time-invariant country characteristics. We now ask: how does the expansiveness of higher education within a country impact wage selection?

In the UK, university attendance and completion is comparatively rare. However, in the last 30 years, the UK has expanded their university system, most notably with the transformation of former polytechnic colleges and central institutions into full-scale universities in 1992, increasing the number of universities from 56 to 94. Even before 1992, attendance at universities was increasing. For example, approximately 20% of the NCDS cohort completed university, most attending between 1976 and 1979. Among the children born in 1970 from the British Cohort Study (BCS)—designed to mirror the NCDS survey—more than 30% completed university, most attending right before the major expansion in 1992. While education expanded, university education remained free until 1998.

Hence a comparison of these two cohorts—the NCDS cohort born in 1958 and the BCS cohort born in 1970—provide a test of the impact of educational expansion on selection into and through university. On the one hand, if the norms, values, and resources associated with higher socioeconomic status are the main predictors of university attendance, rather than individual ability, we would expect that as the educational system expands, opportunities should open up to individuals of lower socioeconomic status and negative selection should decline (H2B). On the other hand, if the most capable individuals are also the most likely to attend and complete university, then expanding the system will increase negative selection by allowing those less capable of capitalizing on a university degree to obtain one, leading to more negative selection (H2A). By comparing a birth cohort that enjoyed greater access to university to a cohort with more restricted access, while cost remained constant, we hope to further illustrate the role that university selectivity plays in selection.

[INSERT FIGURE 5 ABOUT HERE]

Figure 5 shows results of the propensity score analysis from the 1958 and 1970 cohorts in the UK at ages 33 and 34 respectively, alongside the earlier finding from 1958 cohort at age 46, in order to match period of observation as well as age. The graph on the left shows the results from the NCDS cohort born in 1958. Even at age 33, there is a strongly negative trend in returns to a university degree across propensity score strata. Moving from one propensity score stratum to the next decreases one's wage returns 6%. As shown in earlier analyses and repeated here, this negative selection carried forward until mid-career as well. Due to educational expansion in the 1970s and 1980s, the BCS cohort experienced greater opportunity to attend university. Their wage returns are shown on the right side of figure 5. Although still negative, the linear trend of their wage returns is less steep and no longer statistically significant. Moving from one stratum to the next decreases wage returns just over 1%, compared to others within their propensity score strata. Although we cannot test directly that this changing selectivity is the result of educational expansion, these results provide suggestive evidence that expansion helps to get those who would benefit most from completing university to actually attend and complete.

Discussion

Brand and Xie's work on returns to college in the United States presents a compelling behavioral model of negative economic selection into University (Hout 2012) yet this model and the propensity score methodology employed is still strongly debated by those who espouse the hypothesis of positive selection into university completion (Heckman et al. 2006; Carneiro et al. 2011; Zhou and Xie 2016). Moreover the US is an exceptional case of higher education in many ways. Our article therefore develops this important work further, applying the propensity score matching model to new contexts and drawing on the comparative stratification literature to formulate hypotheses of institutional variation underlying population level heterogeneity in the economic returns to higher education.

We first replicate Brand and Xie for the US, demonstrating as did they that there is no statistically significant economic selection in university completion in the USA. We then go on to link three institutional factors with economic selection into higher education: the expansiveness of the university sector, non-academic post-secondary training options, and educational costs. In general, the economic model of university selection anticipates that more selective, binary systems should create more efficient economic sorting and hence more positive selection. In contrast, non-economic models of university completion anticipate that more expansive systems should reduce the importance of prestige in university decision making and result in more positive economic selection. As applied to the three case studies here, economic hypotheses would expect more positive selection in the UK and Germany than the US, and non-economic models less positive selection.

Overall we only find evidence of economic selection into university in the UK. In the US and Germany, we find slopes in the wage returns across propensity strata that do not differ significantly from 0, suggesting no selection at the population level. To explore this surprising null finding further, we conduct extensive robustness checks, and continue to find no linear relationship between propensity to complete university and wage returns in mid-career in these two countries.

We next investigated the effect of a particular institutional factor – a lack of educational expansion - might explain the negative selection in the UK. To do so, we examined changes in selection across two points in the UK, during a time of substantial expansion but constant tuition costs, comparing the selection observed among men of the 1958 birth cohort with those of the 1970 birth cohort. We find that the negative selection observed prior to higher education expansion disappears when we look at selection in the expanded higher education system encountered by the later cohort. This provides support for hypothesis H2B and suggests that educational expansion might explain the lack of negative selection observed in the USA, yet documented in earlier cohorts of the UK. However, our analysis cannot shed any light on why we do not observe selection into university completion in Germany. Auxiliary analyses comparing selection differences into academic versus vocational post-secondary schooling in Germany was suggestive of more negative selection into university rather than vocational tracks,[1] but the differences were not statistically significant and so we do not report these results here.

In sum, our test for population level heterogeneity in the economic returns to university completion, using the propensity score matching method employed by Brand and Xie, does not offer empirical support for either the positive or negative selection hypotheses. However, it is important to note that the majority of our analyses found no statistically significant evidence of a linear trend in the returns to a degree by propensity to complete. It is difficult to definitively explain our null finding. On the one hand, it is possible that the assumptions of our models are upheld and there is no actual linear trend in the economic returns to university by propensity to complete in the United States and Germany. If this is the case, then we can conclude that

[1] Available from authors on request.

negative selection is occurring only in an institutional setting defined by very low financial and opportunity costs but relatively meager vocational training offerings and expansion.

On the other hand, it is possible that there is population level heterogeneity, but that our data is not sufficiently rich to adequately match university graduates with non-graduates, creating unobserved heterogeneity which biases our results. We believe that this is likely to be the case for our analysis of the German data, but not for the USA and the UK. Propensity score matching relies on large samples of both treated and untreated respondents, as well as very rich data which reduces the possibility of unobserved variable bias. The UK data best fulfils these requirements, offering the most complete array of socioeconomic, social-psychological, and cognitive ability indicators, measured prior to treatment, and including a very large sample size. Although the US dataset is smaller, university completion was a less rare event in this context and therefore there is ample sample sizes of completers and non-completers across strata; moreover the US data shares with the UK data a set of rich set prospective measures, including a cognitive ability measure. We therefore feel more confident in the null results we report for the US, and the negative but improving selection we report across the UK cohorts.

Despite these limitations, this study presents a first examination of cross-national variation in selectivity into university completion, and develops for the first time a series of hypotheses linking institutional differences in higher education to economic sorting into university completion. Unlike previous work, our analysis finds little support for the general predictions of negative and positive selection hypotheses. However, we show that selection can and does vary across space (USA and UK) and time (cohorts within the UK) and thus is potentially amenable to policy change. Moreover we also provide further evidence of the very stringent data demands of the propensity score matching method of distinguishing heterogeneous effects (Breen, Choi and Holm 2015), demonstrating that it is unsuitable for smaller samples lacking cognitive ability measures, such as those in our German data.

Popular press and academic publications alike have focused on the possible detrimental effect of higher tuition costs in the US, the UK, and Germany for discouraging university completion by less advantaged youth, yet the results of our analysis suggest that cost is a less important predictor of selection in university completion than more open entrance policies. The least expensive university option in our comparison was the UK, and this was the only sample for whom we observed negative selection. Yet, after undergoing a period of educational expansion, later UK cohorts did not demonstrate the same negative selection patterns. It would be intriguing to conduct further cross-temporal comparisons, for instance following rising tuition costs in these countries or during the period of still further expansion, alongside increased marketization, which has occurred in the past 15 years in the UK. In sum, more comparative research, with increasingly rich datasets, should enable the identification of more of the institutional mechanisms behind selection in higher education in the future.

REFERENCES

Ainsworth, J. W., & Roscigno, V. J. (2005). Stratification, school-work linkages and vocational education. *Social Forces, 84*(1), 257-284.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*: University of Chicago Press.

Bar Haim, E., & Shavit, Y. (2013). Expansion and inequality of educational opportunity: A comparative study. *Research in Social Stratification and Mobility, 31*(0), 22-31, doi:http://dx.doi.org/10.1016/j.rssm.2012.10.001.

Beattie, I. R. (2002). Are all" adolescent econometricians" created equal? Racial, class, and gender differences in college enrollment. *Sociology of Education*, 19-43.

Becker, G. S. (1994). *Human capital: A theoretical and empirical analysis, with special reference to education*: University of Chicago Press.

Becker, R., & Hecken, A. E. (2009). Why are Working-class Children Diverted from Universities?—An Empirical Assessment of the Diversion Thesis. *European Sociological Review, 25*(2), 233-250, doi:10.1093/esr/jcn039.

Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *Stata Journal, 2*(4), 358-377.

Blossfeld, H.-P., & Stockmann, R. (1999). The German Dual System in Comparative Perspective. *International Journal of Sociology, 28*, 3-28.

Blossfeld, H.-P., & Timm, A. (2003). *Who marries whom?: educational systems as marriage markets in modern societies* (Vol. 12): Springer Science & Business Media.

Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in education, society and culture* (Vol. 4). London: Sage.

Brand, J. E., & Davis, D. (2011). The impact of college education on fertility: Evidence for heterogeneous effects. *Demography, 48*(3), 863-887.

Brand, J. E., & Xie, Y. (2010). Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education. *American Sociological Review, 75*(2), 273-302, doi:Doi 10.1177/0003122410363567.

Breen, R., Choi, S., & Holm, A. (2015). Heterogeneous Causal Effects and Sample Selection Bias. *Sociological Science, 2*, 351-369, doi:10.15195/v2.a17.

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and society, 9*(3), 275-305.

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*, 319-337.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics, 3*, 1801-1863.

Card, D., & Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics, 116*(2), 705-746, doi:10.1162/00335530151144140.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating Marginal Returns to Education. *American Economic Review, 101*(6), 2754-2781, doi:doi: 10.1257/aer.101.6.2754.

Carnevale, A. P., Rose, S. J., & Cheah, B. (2011). The College Payoff: Education, Occupations, Lifetime Earnings. *Georgetown University Center on Education and the Workforce*.

Cheung, S. Y., & Egerton, M. (2007). Great Britain: Higher Education Expansion and Reform - Changing Educational Inequalities. In Y. Shavit, R. ARum, & A. Gamoran (Eds.), *Stratification in higher education: A comparative study*: Stanford Univ Press.

Dearden, L., McGranahan, L., & Sianesi, B. (2004). Returns to education for the'marginal learner': evidence from the BCS70. *Centre for the Economics of Education, London School of Economics and Political Science*,

Deil-Amen, R., & Turley, R. L. (2007). A review of the transition to college literature in sociology. *Teachers College Record, 109*(10), 2324-2366.

Dex, S. (1995). The Reliability of Recall Data: A Literature Review. *Bulletin de Methodologie Sociologique, 49*, 58-89.

Douglas, J. W. B. (1964). *The home and the school; a study of ability and attainment in the primary school* (Studies in society). London,: Macgibbon & Kee.

Gangl, M. (2010). Causal Inference in Sociological Research. *Annual Review of Sociology, 36*(1), 21-47, doi:doi:10.1146/annurev.soc.012809.102702.

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research, 25*(3), 201-239.

Goebel, J. (Ed.). (2015). *Documentation on Biography and Life History Data for SOEP v30* (SOEP Survey Papers 266: Series D). Berlin: DIW/SOEP.

Grodsky, E., & Jones, M. T. (2007). Real and imagined barriers to college entry: Perceptions of cost. [doi: 10.1016/j.ssresearch.2006.05.001]. *Social Science Research, 36*(2), 745-766.

Hanushek, E. A., & Wößmann, L. (2006). Does Educational Tracking Affect Performance and Inequality? Differences- in-Differences Evidence Across Countries*. *The Economic Journal, 116*(510), C63-C76, doi:10.1111/j.1468-0297.2006.01076.x.

Hauser, R. M., Tsai, S. L., & Sewell, W. H. (1983). A model of stratification with response error in social and psychological variables. *Sociology of Education, 56*(1), 20-46.

Heckman, J. J., Urzua, S., & Vytlacil, E. J. (2006). Understanding instrumental variables in models with essential heterogeneity. National Bureau of Economic Research.

Heller, D. E. (1997). Student price response in higher education: An update to Leslie and Brinkman. *Journal of Higher education*, 624-659.

Henderson, D. J., Polachek, S. W., & Wang, L. (2011). Heterogeneity in schooling rates of return. *Economics of Education Review, 30*(6), 1202-1214, doi:http://dx.doi.org/10.1016/j.econedurev.2011.05.002.

Hillmert, S., & Jacob, M. (2003). Social Inequality in Higher Education. Is Vocational Training a Pathway Leading to or Away from University? *European Sociological Review, 19*(3), 319-334, doi:10.1093/esr/19.3.319.

Hilmer, M. J., & Hilmer, C. E. (2012). On the relationship between student tastes and motivations, higher education decisions, and annual earnings. *Economics of Education Review, 31*(1), 66-75, doi:http://dx.doi.org/10.1016/j.econedurev.2011.09.004.

Hout, M. (2012). Social and Economic Returns to College Education in the United States. *Annual Review of Sociology, 38*(1), null, doi:doi:10.1146/annurev.soc.012809.102503.

Hunter, J. E., Crosson, J. J., & Friedman, D. H. (1985). The validity of the Armed Services Vocational Aptitude Battery (ASVAB) for civilian and military job performance. Washington, DC: Department of Defense.

Kalleberg, A. L. (2000). Nonstandard employment relations: Part-time, temporary and contract work. *Annual Review of Sociology, 26*, 341-365.

Lau, L. K. (2003). Institutional factors affecting student retention. *Education, 124*(1).

Long, M. C. (2010). Changes in the returns to education and college quality. *Economics of Education Review, 29*(3), 338-347.

Lucas, S. R. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects1. *American Journal of Sociology, 106*(6), 1642-1690.

Lunt, M. (2013). pbalchk: Stata module for checking the balancing of the covariates between two groups.

Malamud, O., & Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics, 95*(11–12), 1538-1549, doi:http://dx.doi.org/10.1016/j.jpubeco.2011.03.006.

McLanahan, S., & Sandefur, G. D. (1994). *Growing up with a single parent: What hurts, what helps*: Harvard University Press.

Meek, V. L., Goedegebuure, L., Kivinen, O., & Rinne, R. (1996). *The Mockers and Mocked: Comparative Perspectives on Differentiation, Convergence, and Diversity in Higher Education*: IAU Press Oxford.

OECD (2011). Education at a Glance 2011. OECD.

Paterson, L., & Iannelli, C. (2007). Social Class and Educational Attainment: A Comparative Study of England, Wales, and Scotland. *Sociology of Education, 80*(4), 330-358, doi:10.1177/003804070708000403.

Perna, L. W., & Titus, M. A. (2005). The Relationship between Parental Involvement as Social Capital and College Enrollment: An Examination of Racial/Ethnic Group Differences. *The Journal of Higher Education, 76*(5), 485-518.

Quinn, J. (2013). Drop-out and completion in higher education in Europe among students from under-represented groups. *Network of Experts on Social aspects of Education and Training (NESET)*. European Union: European Commission.

Raftery, A. E., & Hout, M. (1993). Maximally maintained inequality: Expansion, reform, and opportunity in Irish education, 1921-75. *Sociology of Education*, 41-62.

Raftery, A. E., & Hout, M. (1993). Maximally maintained inequality: Expansion, reform, and opportunity in Irish education, 1921-75. *Sociology of Education, 66*(1), 41-62.

Shavit, Y., Arum, R., Gamoran, A., & Menachem, G. (2007). *Stratification in higher education: A comparative study*: Stanford Univ Pr.

Shavit, Y., & Muller, W. (2000). Vocational Secondary Education. *European Societies, 2*, 29-50.

Shavit, Y., & Müller, W. (1998). *From school to work: a comparative study of educational qualifications and occupational destinations*. Oxford: Clarendon Press

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355-374.

Tam, T., & Jiang, J. (2014). The Making of Higher Education Inequality: How Do Mechanisms and Pathways Depend on Competition? *American Sociological Review, 79*(4), 807-816, doi:10.1177/0003122414534437.

Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective. In K. S. Cook, & D. S. Massey (Eds.), *Annual Review of Sociology, Vol 36* (Vol. 36, pp. 407-428, Annual Review of Sociology). Palo Alto: Annual Reviews.

Willis, R. J., & Rosen, S. (1978). Education and self-selection. National Bureau of Economic Research Cambridge, Mass., USA.

Xie, Y., Brand, J. E., & Jann, B. (2011). Estimating heterogeneous treatment effects with observational data. *Sociological Methodology*.

Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating heterogeneous treatment effects with observational data. *Sociological Methodology, 42*(1), 314-347.

Table 1. Comparison of University Characteristics for Men Born 1955-1965

	Germany	United Kingdom	United States
Stucture[a]			
Binary/Diverse/Unified	Binary	Binary	Diverse
Eligibility and Enrollment[b]			
% men ages 45-54 eligible for higher education (*Hochschulreife,* 2+ A-levels, Highschool degree)	31%	16%	89%
% men ages 45-54 in 2010 at least university degree	16%	23%	29%
Market structure / Costs[c]			
Low/Moderate/ High Private Funding	Low	Low	High
Opportunity Cost	$171,969	$98,358	Public: $230,308 Private: $259,308

a) Binary, Diverse and Unified education systems as defined in Shavit et al (2007). Unified systems offer primarily academic tertiary education, binary systems offer a second tier of vocational training, and diverse systems include both academic and vocational training in a second tier

b) Expansion estimates, UK: Machin and McNally 2006; Germany: Statistisches Bundesamt 2014; USA: OECD 2012;

c) Cost estimates: Day and Newburger 2002

Table 2. Mean values of analysis variables by non-university/university graduate status and survey

	UK (NCDS)		Germany GSOEP 2000		Germany GSOEP 2004		USA (NLSY)	
	Non-Grad	Grad	Non-Grad	Grad	Non-Grad	Grad	Non-Grad	Grad
Dependent Variable								
Gross Monthly Wages in 2000 Euros								
2000, Age 42	3574.68	5315.03						
2004, Age 46	4217.27	6327.49						
2000, Ages 35-45			2892.41	3769.11				
2004, Ages 39-49					3265.57	4405.16		
Hourly wages and salary in dollars								
2000, Ages 35-38							15.88	26.85
2004, Ages 39-42							16.71	33.86
Variables Measuring Propensity to Complete University								
Social Background								
Father's ISEI score			39.71	53.12	39.35	50.68		
Father's Class: Upper	.18	.43						
Father's Class: Middle	.53	.42						
Father's Class: Lower	.25	.11						
Father's Years Education	10.07	11.45	11.53	12.93	11.55	12.74	11.44	14.21
Mother's Years Education	9.81	10.68	10.33	11.49	10.39	11.40	11.40	13.25
Family income	289.83	298.03					17878.26	24782.01
Rural	.50	.52	.61	.57	.60	.56	.25	.20
Intact family	.74	.76	.90	.91	.87	.88	.72	.82
Siblings	2.10	1.87	1.60	1.44	1.82	1.60	3.23	2.39
Proximity to college							.79	.77
Ethno-religious								
Black							.18	.08
Hispanic							.07	.04
Jewish							.00	.03
Mother Catholic			.43	.40	.44	.40		
Parent foreign born	.09	.12						
Ability and Academics								
Grades / Exam scores	.26	.58	3.70	4.12	3.79	4.09		
Had an Apprenticeship			.93	.29	.91	.25		
Abitur / Alevels / College prep	.07	.72	.19	.97	.20	.96	.24	.60
Math ability age 16	2.80	1.57						
English ability age 16	2.66	1.59						
Mental ability							.05	.79
Social-Psychological								
Parents care / Help school	.56	.56	2.42	2.41	2.64	2.73		
Peers plan attend college	.06	.03					.42	.80
R's attitude towards school at 16 - index score	3.64	4.10						

Note: Employed men with non-zero wages, self-employed excluded. Grades (for Germany) are an average of retrospective reports of the final grade on report card in secondary school (1-6). Exam scores (for UK) is a summary score from 0-1 of A-levels, O-levels, and CSE exams passed from official administrative reports. Math and English ability at age 16 are reports from school officials about the student's math or english language ability scaled high to low 1-6

Table 3 Homogenous effects of a university degree on logged wages

	USA:NLSY						UK:NCDS						Ger	
	2000			2004			2000			2004			2000	
	35-38			39-42			Age 42			Age 46			Ages 35-45	
Homogenous Returns														
	Coef.	SE		Coef.	SE		Coef.	SE		Coef.	SE		Coef.	SE
University Completion	.247	.073	*	.161	.107		.063	.043		.158	.044	*	.137	.057
R^2	.211			.154			.128			.165			.188	
N	838			770			2682			1911			700	

Note: Employed men with non-zero wages, self-employed excluded. * p<.05. All models include the controls listed in table 2.

Table 4. Description of propensity score strata, GSOEP and NCDS

	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG
	Germany 2000								*Germany 2004*						
	0-.4		.4 -.6		.6-.8		.8-1		0-.2		.2-.4		.4-.6		
Father's ISEI Score	38.81	41.26	49.99	50.66	53.91	54.13	49.18	57.60	37.61	35.59	44.76	48.61	44.14	45.89	45.
Father's Years Education	11.43	11.87	12.51	12.28	14.00	13.05	11.78	13.47	11.35	11.30	12.23	12.63	11.64	11.96	12.
Mother's Years Education	10.25	10.91	11.73	11.03	11.20	10.90	10.44	12.20	10.26	10.80	10.55	11.13	11.27	11.60	11.
Rural	.62	.70	.68	.59	.40	.46	.67	.59	.62	.50	.55	.60	.45	.57	.3
Intact Family	.90	.92	.94	.89	.92	.87	.98	.93	.88	.88	.82	.91	.93	.79	.8
Siblings	1.63	1.17	1.18	1.44	1.60	1.87	1.22	1.29	1.87	2.70	1.82	1.55	1.41	1.39	1.4
Mother Catholic	.42	.41	.40	.48	.46	.38	.73	.38	.45	.24	.36	.60	.37	.38	.5
Grades	3.69	4.00	3.96	4.15	3.69	3.91	3.85	4.28	3.79	3.98	3.72	3.82	4.05	4.09	3.6
Had Apprenticeship	.96	.96	.91	.84	.20	.03	.00	.00	.96	.80	.88	.95	1.00	.87	.1
Had abitur	.13	.78	1.00	1.00	1.00	1.00	1.00	1.00	.02	.30	.88	.95	1.00	1.00	1.0
Parent's care about schooling	2.42	2.29	2.41	2.34	2.40	2.46	2.30	2.45	2.63	2.53	2.62	2.73	2.65	2.59	2.8
Age	39.59	40.26	39.14	39.56	39.50	40.18	42.33	39.83	43.61	43.40	43.13	43.40	44.18	44.43	44.
Married 2004	.77	.74	.77	.80	.80	.66	.89	.73	.77	.80	.68	.85	.73	.83	.9
Has child 2004	.71	.61	.68	.72	.70	.58	.78	.67	.65	.80	.68	.85	.68	.52	.8
N	507	23	22	25	10	38	9	66	397	10	56	20	22	23	1

Table 4. Description of propensity score strata, GSOEP and NCDS (con'd)

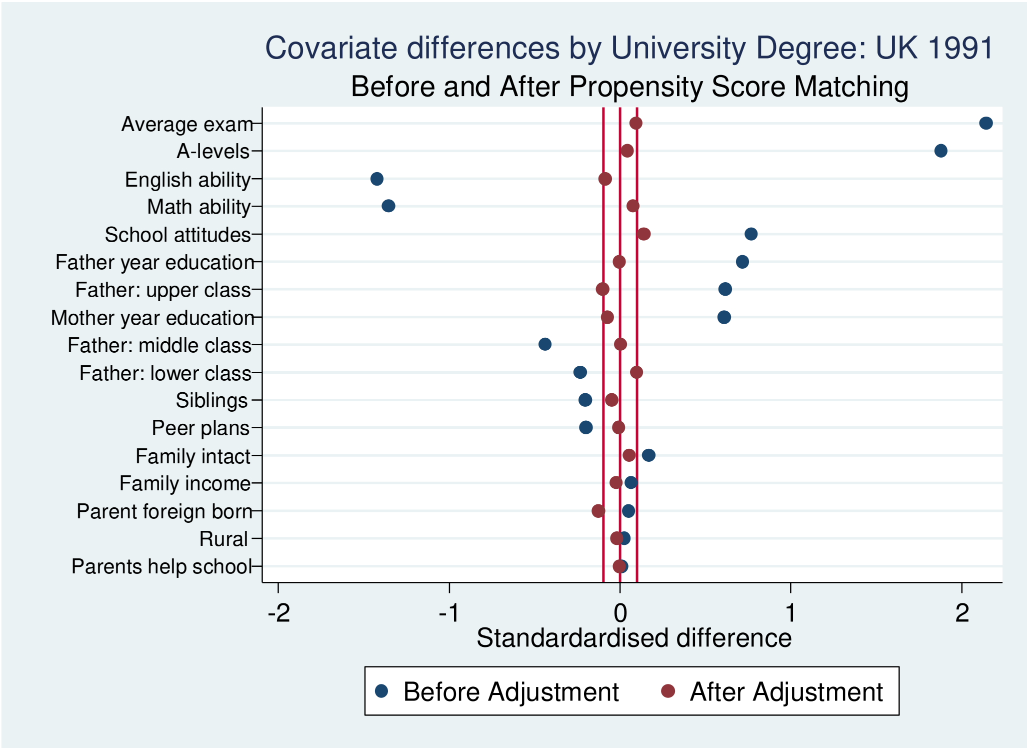
	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG	Grad	NG
	United Kingdom 2000												
	0-.0.02		0.35-0.05		0.05-.1		.1-.2		.2-.4		.4-.6		.
Father's Class: Upper	.08	.11	.17	.18	.19	.29	.30	.27	.43	.28	.34	.35	.44
Father's Class: Middle	.54	.56	.55	.45	.56	.57	.52	.46	.46	.57	.43	.41	.44
Father's Class: Lower	.34	.33	.25	.36	.21	.07	.14	.22	.09	.13	.14	.16	.12
Father's Years Education	9.71	9.78	9.92	9.66	10.11	9.82	10.55	10.55	10.94	10.96	10.77	11.04	11.08
Mother's Years Education	9.56	10.22	9.71	9.64	9.95	9.90	9.99	9.89	10.41	10.20	10.46	10.21	10.61
weekly family income age 16	278	304	296	266	300	275	302	337	295	319	293	291	291
Rural	.52	.59	.49	.36	.49	.50	.43	.60	.59	.50	.45	.53	.54
Intact family	.71	.78	.74	.82	.77	.68	.76	.68	.75	.79	.82	.79	.78
Siblings	2.23	2.89	2.21	2.00	1.90	1.82	2.01	1.99	1.76	1.52	1.88	1.80	1.83
Parent foreign born	.07	.11	.10	.18	.10	.00	.09	.22	.12	.16	.12	.10	.06
Exam scores	.17	.17	.22	.25	.27	.29	.34	.34	.45	.45	.55	.58	.61
A-levels	.00	.00	.00	.07	.00	.01	.01	.00	.16	.15	.74	.82	.96
Math ability 16	3.45	3.70	2.80	2.58	2.41	2.22	2.10	1.94	1.82	1.76	1.83	1.88	1.44
English ability 16	3.22	2.89	2.68	2.48	2.33	2.16	2.06	1.96	1.85	1.69	1.64	1.76	1.57
Parents help with school	.55	.44	.57	.55	.54	.53	.56	.53	.58	.49	.57	.55	.57
Peers plan to attend college	.08	.22	.05	.00	.03	.00	.05	.04	.02	.02	.05	.02	.02
R's school attitude	3.29	3.35	3.77	3.85	3.89	3.91	3.99	3.94	4.04	4.11	3.87	4.03	4.01
Child in HH - 2000	.53	.56	.53	.91	.60	.54	.61	.70	.68	.66	.69	.60	.62
Married in 2000	.72	.78	.74	.82	.77	.75	.78	.78	.75	.75	.74	.71	.82
N	892	9	412	11	359	28	236	37	127	53	70	68	50

Table 5. Heterogeneous effects of a university degree on logged wages

	United States: NLSY						United Kingdom: NCDS						German	
	2000			2004			2000			2004			2000	
	35-38			39-42			Age 42			Age 46			Age 35-45	
Heterogeneous Returns														
University Completion Strata														
1	.402	.243		-.106	.719		.107	.107		.183	.074	*	.224	.084
	[611]			[611]			[901]			[901]			[530]	
2	-.427	.314		.096	.181		.004	.097		.297	.105	*	.158	.084
	[141]			[141]			[423]			[423]			[47]	
3	.432	.085	*	.308	.106	*	.170	.085	*	.362	.089	*	.253	.205
	[192]			[192]			[387]			[387]			[48]	
4	.465	.162	*	.235	.212		-.109	.091		.162	.141		.057	.155
	[126]			[126]			[273]			[273]			[75]	
5	.404	.140	*	.446	.233		.098	.107		.214	.096	*		
	[120]			[120]			[180]			[180]				
6							.160	.112		.030	.106			
							[138]			[138]				
7							.014	.104		.116	.108			
							[222]			[222]				
8							-.433	.194	*	-.312	.182			
							[158]			[158]				
Slope	.041	.059		.096	.087		-.020	.018		-.038	.016	*	-0.05	0.06
Constant	.256	.210		-.033	.287		.124	.082		.327	.071	*	0.27	0.12

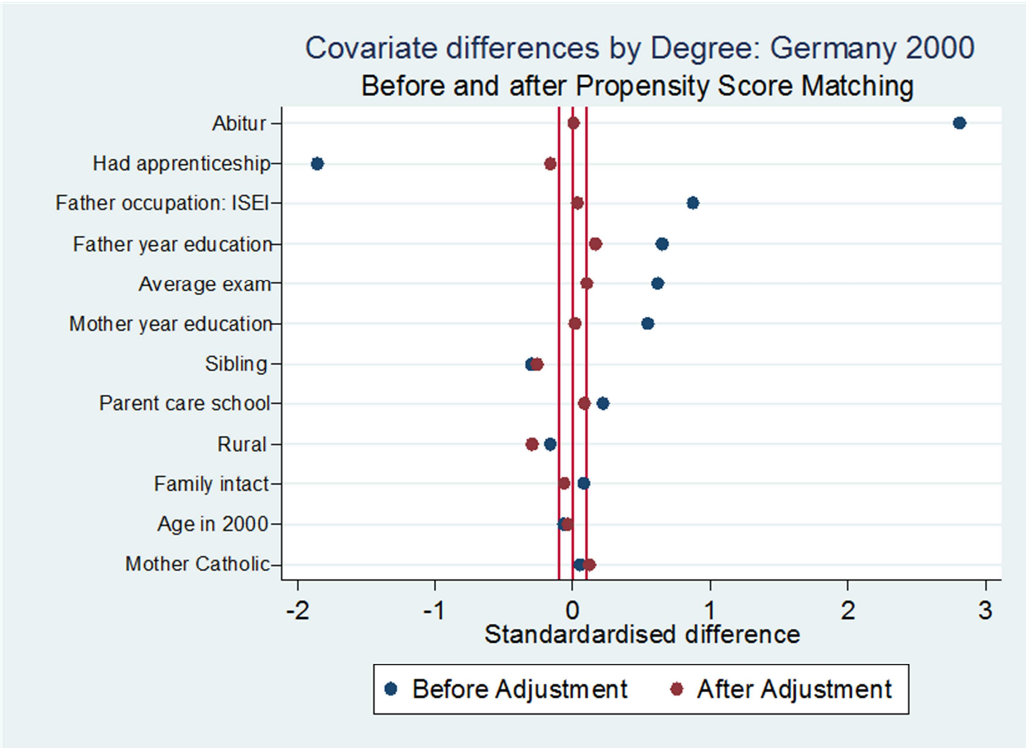
Note: Employed men with non-zero wages, self employed excluded * p<.05.

FIGURE 1



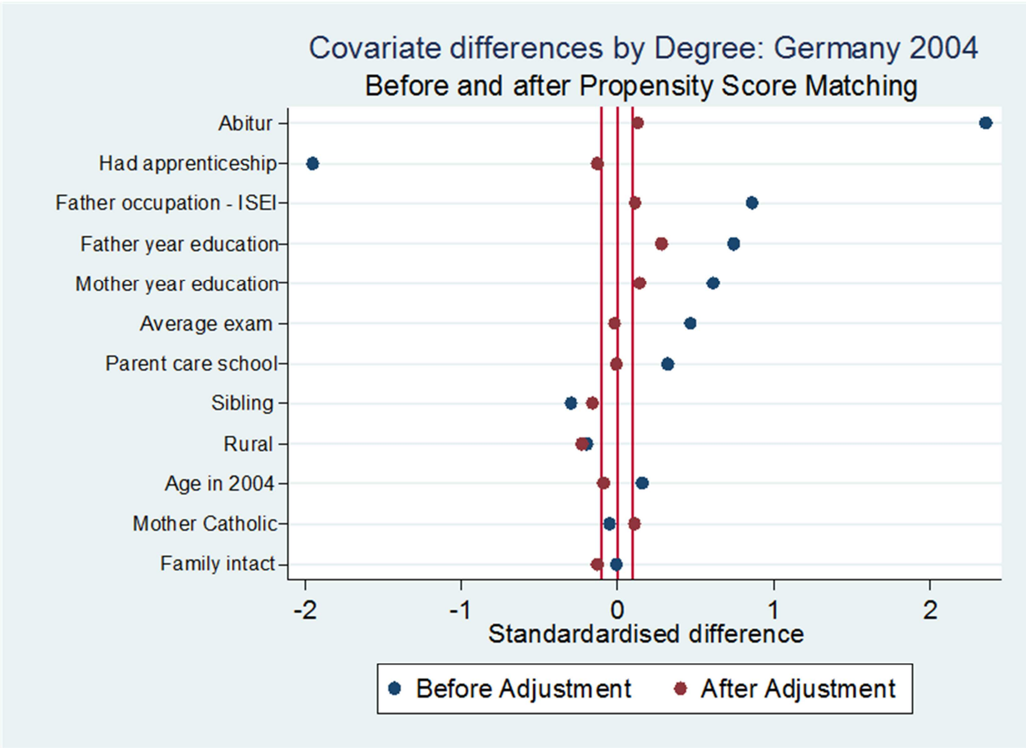
Note: Standardised differences between those with and without a university degree, before and after propensity score matching and strata formation. Red vertical lines denote .01 significance levels.

FIGURE 2



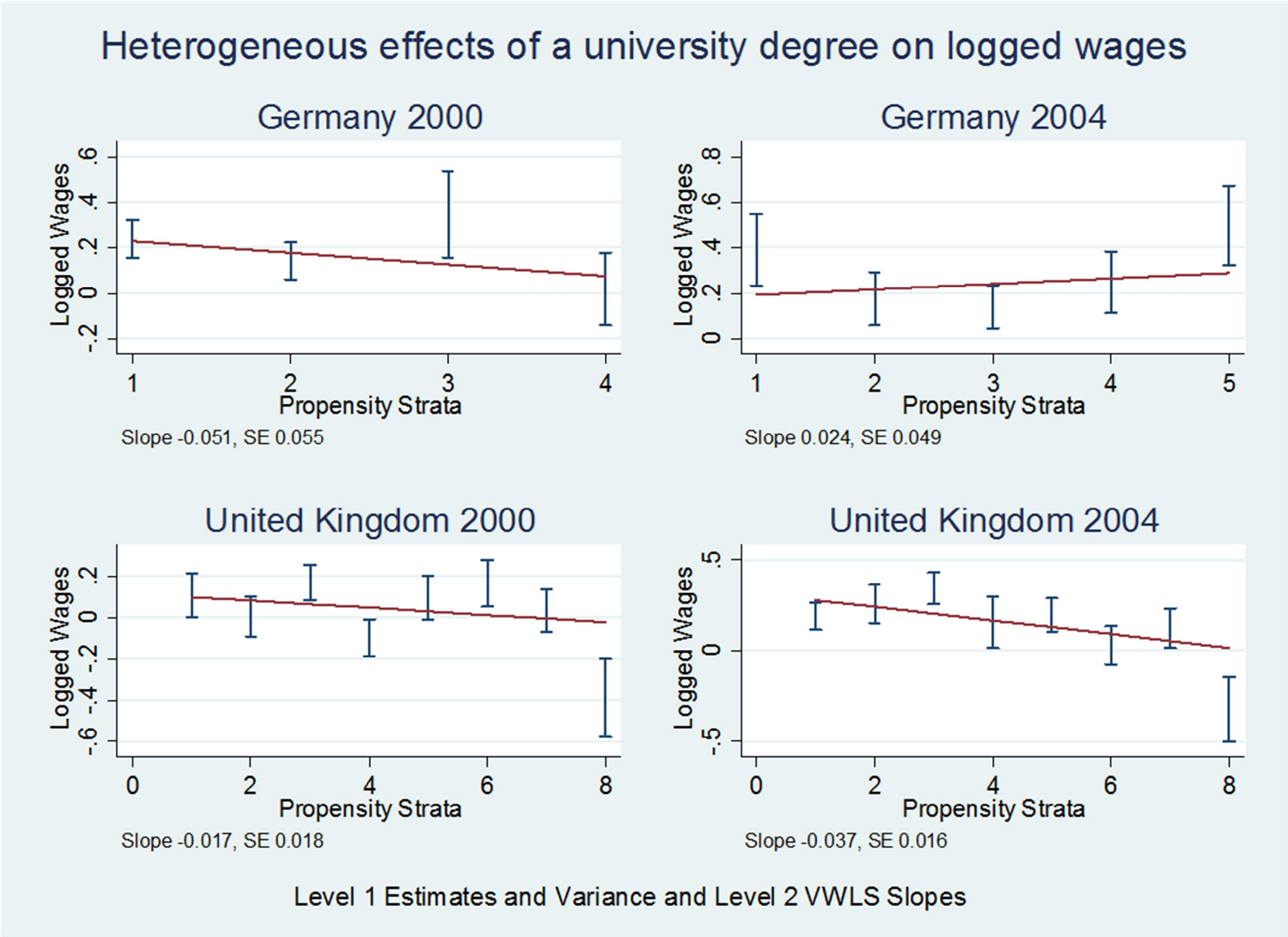
Note: Standardised differences between those with and without a university degree, before and after propensity score matching and strata formation. Red vertical lines denote .01 significance levels.

FIGURE 3



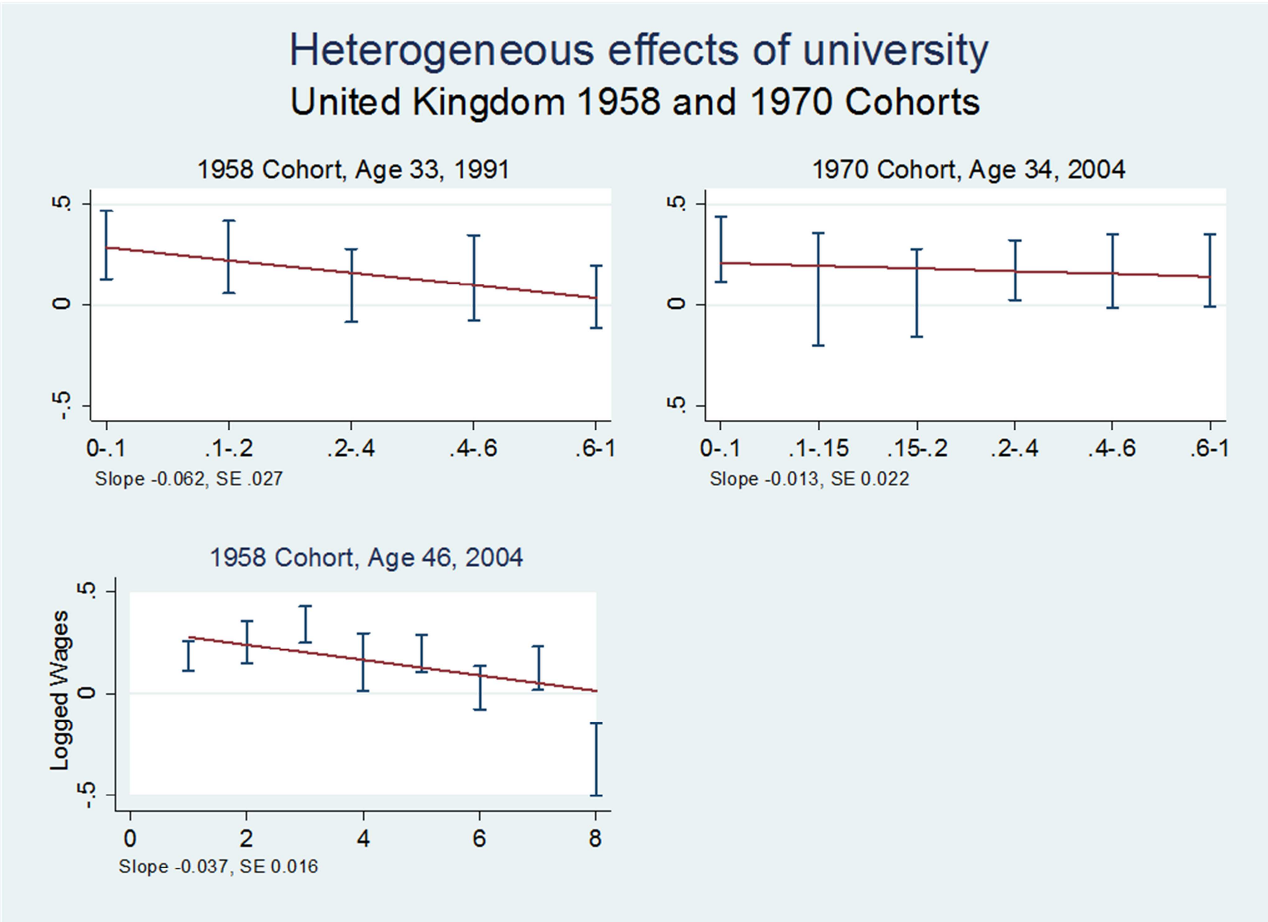
Note: Standardised differences between those with and without a university degree, before and after propensity score matching and strata formation. Red vertical lines denote .01 significance levels.

FIGURE 4



Heterogeneous effects of a university degree on logged wages
Germany 2000
Logged Wages
-.2
0
.2
.4
.6
1
2
3
4
Propensity Strata
Slope -0.051, SE 0.055
Germany 2004
Logged Wages
0
.2
.4
.6
.8
1
2
3
4
5
Propensity Strata
Slope 0.024, SE 0.049
United Kingdom 2000
Logged Wages
-.6
-.4
-.2
0
.2
0
2
4
6
8
Propensity Strata
Slope -0.017, SE 0.018
United Kingdom 2004
Logged Wages
-.5
0
.5
0
2
4
6
8
Propensity Strata
Slope -0.037, SE 0.016
Level 1 Estimates and Variance and Level 2 VWLS Slopes

FIGURE 5



Heterogeneous effects of university
United Kingdom 1958 and 1970 Cohorts
1958 Cohort, Age 33, 1991
.5
0
-.5
0-.1
.1-.2
.2-.4
.4-.6
.6-1
Slope -0.062, SE .027
1970 Cohort, Age 34, 2004
.5
0
-.5
0-.1
.1-.15
.15-.2
.2-.4
.4-.6
.6-1
Slope -0.013, SE 0.022
1958 Cohort, Age 46, 2004
Logged Wages
.5
0
-.5
0
2
4
6
8
Slope -0.037, SE 0.016

APPENDIX A1

Estimated relationship between college completion and log wages under the assumption of homogeneity, NLSY							
	Employed Respondents				B&X Results		
	Coef.	SE	R^2	N	Coef.	SE	N
1994 (29-32)	0.133	0.052	0.20	877	0.180	0.047	862
1996 (31-34)	0.226	0.059	0.22	906			
1998 (33-36)	0.274	0.058	0.22	882	0.296	0.054	881
2000 (35-38)	0.247	0.073	0.21	838			
2002 (37-40)	0.402	0.082	0.23	796	0.410	0.069	777
2004 (39-42)	0.161	0.107	0.15	770			

Effects of college completion on log wages by propensity score strata, NLSY						
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	Level 2 Slopes
1994 (29-32)						
B&X (Employed)	0.351	0.164	0.099	0.212	0.082	-0.046
	(0.123)	(0.121)	(0.107)	(0.104)	(0.082)	(0.038)
Employed	0.416	0.047	0.148	0.149	0.142	-0.012
	(0.229)	(0.142)	(0.087)	(0.129)	(0.097)	(0.044)
1996 (31-34)						
Employed	0.533	0.010	0.274	0.216	0.376	0.007
	(0.173)	(0.124)	(0.091)	(0.105)	(0.163)	(0.048)
1998 (33-36)						
B&X (Employed)	0.542	0.214	0.225	0.479	0.197	-0.019
	(0.150)	(0.108)	(0.099)	(0.120)	(0.164)	(0.046)
Employed	0.462	0.005	0.315	0.510	0.327	0.032
	(0.138)	(0.122)	(0.093)	(0.133)	(0.134)	(0.043)
2000 (35-38)						
Employed	0.402	-0.427	0.432	0.465	0.404	0.041
	(0.243)	(0.314)	(0.085)	(0.162)	(0.140)	(0.059)
2002 (37-40)						
B&X (Employed)	0.74	0.234	0.387	0.479	0.408	-0.053
	(0.163)	(0.191)	(0.133)	(0.188)	(0.178)	(0.055)
Employed	0.672	-0.011	0.456	0.526	0.569	0.041
	(0.260)	(0.271)	(0.112)	(0.201)	(0.157)	(0.065)
2004 (39-42)						
Employed	-0.106	0.096	0.308	0.235	0.446	0.096
	(0.719)	(0.181)	(0.106)	(0.212)	(0.233)	(0.087)

APPENDIX A2

Comparing homogeneous and heterogeneous returns in GSOEP and NEPS						
	GSOEP			NEPS		
	2012			2010		
	Age 47-57			Age 52-58		
Homogenous Returns						
	Coef.	SE		Coef.	SE	
University Completion	0.222	0.082	*	0.108	-0.070	
R^2	0.259			0.215		
N	638			892		
Heterogeneous Returns						
University Completion Strata						
1	0.425	0.245		0.324	0.115	*
2	0.278	0.329		-0.794	0.618	
3	0.205	0.214		0.195	0.159	
4	0.237	0.108	*	0.044	0.115	
5	0.734	0.216	*			
6	0.047	0.166				
7						
8						
Slope	-0.034	0.052		-0.086	0.054	
Constant	0.407	0.220		0.397	0.157	*

Note: Employed men with non-zero wages, self-employed excluded. GSOEP includes the controls listed in table 2. The NEPS model includes controls for migration background, mother's education, father's education, father's occupation, intact family, number of siblings, rural residence, grades in final year of school, whether completed *Abitur,* type of secondary school attended, and a measure of whether the adult respondent values learning.

APPENDIX A3

Heterogeneous effects of a university degree, including controls for marriage and children												
	United Kingdom: NCDS						Germany: GSOEP					
	2000			2004			2000			2004		
	Age 42			Age 46			Age 35-45			Age 39-49		
Heterogeneous Returns												
University Completion Strata												
1	.104	.106		.187	.074	*	.240	.082	*	.392	.157	*
	[901]			[901]			[530]			[407]		
2	.004	.096		.256	.104	*	.142	.081		.175	.118	
	[423]			[423]			[47]			[76]		
3	.169	.085	*	.341	.088	*	.347	.189		.140	.094	
	[387]			[387]			[48]			[45]		
4	-.098	.090		.156	.141		.019	.159		.248	.134	
	[273]			[273]			[75]			[53]		
5	.096	.107		.198	.093	*				.498	.174	*
	[180]			[180]						[104]		
6	.166	.114		.030	.104							
	[138]			[138]								
7	.033	.104		.125	.108							
	[222]			[222]								
8	-.388	.189	*	-.321	.176							
	[158]			[158]								
Slope	-.017	.018		-.037	.016	*	-.051	.055		.024	.049	
Constant	.116	.081		.315	.070	*	.280	.116	*	.169	.153	

Note: Employed men with non-zero wages, self employed excluded * p<.05. All models include controls for marital status and number of children

Homogeneous and heterogeneous effects of a university degree on earnings, including 0 wages and self employed												
	United Kingdom: NCDS						Germany: GSOEP					
	2000			2004			2000			2004		
	Age 42			Age 46			Age 35-45			Age 39-49		
Homogenous Returns												
University Completion	224.16	223.82					369.13	153.41		552.86	186.00	*
R^2	.09						.10			.09		
N	3525						833			863		
Heterogeneous Returns												
University Completion Strata												
1	524.83	508.32		Self employed earnings information missing			613.15	204.47	*	407.89	377.99	
2	909.32	526.47					200.30	351.03		442.88	384.59	
3	641.75	345.43					463.02	482.30		457.31	411.29	
4	-380.23	456.18					756.75	494.01		536.31	480.03	
5	527.51	452.75								586.93	562.47	
6	797.12	466.24										
7	-1198.16	929.66										
8	-958.23	1140.57										
Slope	-122.86	98.88					-10.43	155.46		43.06	144.36	
Constant	871.01	414.69	*				546.78	309.87		355.50	425.16	

Note: All men (including self employed and 0 wages), * p<.05. Homogeneous models include all controls in table 2. All models include controls for marital status and number of children