University education and income – does prior achievement matter?*

Magnus Wikström♣

Christina Wikström♦

November 2011

Abstract

The purpose of this study is to find out if the income premium from university entrance differs with respect to prior achievement as measured by previous grades. Using income at the age of 28 to 30, we analyze if high-achievers have larger income premiums from entering university than low-achievers in a sample of Swedish upper secondary school students. We find that income differences generally are positive, albeit larger for females than for males. It is also found that the income premium is larger for high-achievers than for low-achievers. However, especially for males, the income premium rises only marginally with prior achievement for a large part of the grade distribution, indicating that there are only small differences in the returns to university entrance for a majority of upper secondary school graduates.

JEL classification: I21, J24

Keywords: Income Premium; predictive validity; Upper secondary GPA; Achievement; University entrance

* Many thanks to John Hattie, LizMcKinley, and the Faculty of Education, University of Auckland, for great hospitality during the time for which the first version of the paper was written. Valuable comments from seminars at the Department of Economics and the Department of Applied Educational Science, Umeå University, are appreciated. A research grant from IFAU is gratefully acknowledged.

♣ Corresponding author. Address: Department of Economics, Umeå University, SE-901 87 Umeå, Sweden. Email: magnus.wikstrom@econ.umu.se

♦ Address: Department of Applied Educational Science, Umeå University, SE-901 87 Umeå, Sweden. Email: christina.wikstrom@edusci.umu.se

1. Introduction

There are different underlying principles for how the selection into higher education is organised. Wolming (1999) describes three main approaches that admissions systems usually are based on, either separately or in combination. One approach is egalitarian. Here all eligible applicants are selected or a system of lottery is used, where all eligible applicants receive the same chance of being selected. A second approach is when the selection of applicants is based on a utilitarian principle. According to Epicurus' rule of utility it is the approach that brings the greatest happiness to the greatest number of people. In this context the focus of utility may be on all the stakeholders and not necessarily only the students. The third (and most common approach) is meritocracy. In a meritocratic approach the applicants are ranked by previous merits manifested in grades or test scores, and selected accordingly. The assumptions are that this approach is perceived as transparent and fair, but also that the students who are high performing will perform better in higher education than students with lower performances. If the correlation between instrument and the criteria for study success is strong, the instrument is said to have high predictive validity or predictive strength.

In many systems, upper secondary grades are used as indicators of relevant merits for higher education, and therefore used both for eligibility purposes and instruments for ranking and selecting the students. Combining these two purposes is not without problems, however, since the grades are designed to give information about what a student know and can do at the time when the grade is given, but not necessarily what the student can do in the future. There are also expectations of rank ordering – the higher the grade or test score, the better will the students perform – but also of discrimination between those who will "succeed" and those who will not (see Figure 1 below).

Figure 1 HERE

The quality and characteristics of the selection instrument are important for several reasons. The stakes are high for the applicants – they naturally want to have a fair chance of being admitted but not necessarily be admitted if it turns out they did not have the required previous knowledge to succeed or benefit from the education. There are also a number of potential consequences for other stakeholders if incorrect decisions are made. The quality of the selection instruments is usually determined through its validity, e.g. the instrument's ability to give the expected information without leading to unintended consequences (Messick, 1989). In selection to higher education, its predictive validity is of course of special interest (Lyrén, 2008; Wolming, 1999). To investigate the predictive validity of an instrument, the criterion measure is important. In studies of higher education, the criterion is often some kind of performance measure, in the form of grades or credits.

The principle for validation studies is simple – the predictive strength is the correlation between instrument and criterion measure. However, in practice, such studies are often problematic. First of all there is often more than one instrument. For instance, in the Swedish system there are two selection instruments to higher education, the grade point average (GPA) from upper secondary school and an optional admission test (the SweSAT). The GPA and the SweSAT are not combined, however, as is the case in most other systems where there are two instruments, but used separately. Applicants who have both will be put in two admissions groups and admitted from the group where he or she is ranked the highest. Universities must admit at least thirty percent from each admissions group. Selection effects make it difficult to perform validation studies, as practically all applicants have upper secondary school grades but not all have taken the test.

The criterion can also be problematic, for instance when traditional indicators for academic performance, such as grades or graduation are used, and scales, grading practices or standards varies within and between programmes. This is the case for Sweden; there is no common scale at the university level, but in many courses and programmes the students are graded with either a fail or a pass, which means that there will be little or no variation in the criterion. A common solution is to use university credits: a student is regarded successful if he/she takes the expected number of credits per term or year, and the more credits, the better.[1] One problem with this approach is that it is not necessarily the successful students who take the larger number of credits. A student who is attractive on the job market may leave university education before the student who is less attractive or unable to find a job and hence decides to stay on and take more courses. Furthermore, it is also problematic that missing credits due to study breaks caused by childbirth, illness or sabbaticals will be incorrectly interpreted as non-success.

[1] There is a fairly large number of Swedish studies that address the predictive validity of one or both of the selection instruments mentioned in the text (e.g. Henriksson and Wolming, 1998; Svensson, Gustafsson and Reuterberg, 2001; and Svensson and Nielsen, 2004). See also Wolming & Wikström, 2010 for an overview. Internationally, there is an extensive literature dealing with predicting university grades by high school GPA or SAT scores, see e.g. Cohn et al. (2004) and the references therein.

Another shortcoming that characterizes prediction studies is that the criterion (study success) only can be observed for those who enter higher education (correct prediction and false positive in figure 1). This means that the predictive strength of an instrument is determined solely on a within-group comparison of students admitted. This means that prediction studies are rooted in a meritocratic system. If one opens up for other possible guiding principles of how to organize selection, such as the utilitarian, statements about counterfactual outcomes may become important. In a utilitarian system, we may interpret predictive strength in relation to what would have been had the student not been admitted. In such a comparison, a high achieving student obtaining a larger number of credits/higher GPA should not necessarily be admitted ahead of, say, a middle-achiever since the former may have a better outcome under non-admittance. This kind of interpretation is related to the economic approach of studying the "returns" to education.[2] From society's point of view, students with a large expected income premium from education (or rather admission) may be more beneficial to promote into university than students with a low expected premium. Consequently, there is reason to consider alternatives to traditional prediction studies, to avoid the limitations discussed above and also to revisit the definition of the criterion "academic success".

In this paper, an alternative approach to the traditional prediction studies is explored. Since university education is expected to lead to future employment, and successful students should be attractive on the job market, employment and earnings should be relevant indicators of study success. Hence, the main aim of this paper is to study the relationship between how students fare on the selection instrument and how successful they are on the job market. Since earnings are strongly related to educational outcomes, it is reasonable to use earnings as an indicator of academic success. Consequently, our assumption that income is a criterion for academic success may be an alternative if direct performance (in terms of grades) in higher education is difficult to observe. Moreover, earnings are available for all students that have entered the labour market, making it possible to compare individuals entering higher education with similar individuals that do not enter with respect to the selection criteria. In this paper we will focus on the upper secondary GPA as selection instrument to higher education. Earnings measured at labour market entrance are used as the criterion. Our

[2] The return to education usually applies to completed education, and not, as in this paper, entrance into education. See Mincer (1974) for the basic interpretations of the human capital model and the returns to education and Psacharopoulos and Patrinos (2004) for empirical results from a large number of studies connected to this literature.

measure of academic success is the income premium of admission into university. The empirical analysis uses data including Swedish students born 1972 to 1974 that graduated from a theoretically oriented upper secondary school program. A prior achievement interaction effect is estimated, with the purpose of studying if the income premium differs conditional on differences in the selection instrument.

Apart from extending the previous literature by using earnings as an outcome variable and studying the income premium from entrance, the paper also adds to the literature in other ways. First, we characterize the prior achievement interaction effect using a theoretical model, where the main differences between the criterion used in the paper and the previously used criterion can be discussed. Second, interpreting the income premium from university admission is not without problems. We therefore discuss sources of misinterpretation that have to do with identification of the prior achievement interaction effect. Third, the paper also connects to the literature of successful educational policy from a macro-oriented perspective, and in particular the literature on the expansion of the educational system.[3] Although our paper does not study the development of the return to education during a period of university expansion, the empirical part of the paper concerns a period when the number of enrolees into higher education greatly increased in Sweden. As such it is interesting to see how a relatively large number of weak students (in terms of prior school grades) fare relative to other entrants.

The rest of the paper is organized as follows: in section two a theoretical model is developed. The main purpose is to describe the differences between the labour market information approach and the traditional approach that uses grade points or credits as criteria. We show how these measures are connected, and we also discuss problems interpreting the resulting association (between school grades and income premium) because of selection on unobservable characteristics that we cannot fully control for. Section three describes the data and the methods used in the paper. We propose several different ways of how to estimate the prior achievement interaction effect. The empirical application is based on three cohorts of students on academic track in upper secondary school entering higher education relatively soon after high school completion. In section four the results are presented, and in section five some concluding remarks are made.

[3] See Osborne (2003) for a study on the expansion of higher education in Europe. Walker and Zhu (2008) study the development of college wage premium in Great Britain.

2. A model for interpreting the prior achievement interaction effect

We assume that the observed earnings of an individual reflect factors related to education; $y = y(a, x)$, where y denotes earnings, a achievement at university, and x prior achievement (corresponding to upper secondary school or high-school). Earnings are assumed to be increasing and strictly concave in both arguments, meaning that the labour market values both types of achievement but at a decreasing rate. Let super indices e and 0 denote entrants and non-entrants respectively. The earnings difference between a university entrant and a non-entrant, for a given prior achievement level, can be expressed as $\Delta y = y^e - y^0$. The main concern here is to evaluate how the income difference between entrants and non-entrants depends on prior achievement;

$$\frac{\partial \Delta y}{\partial x} = \frac{\partial y^e}{\partial x} - \frac{\partial y^0}{\partial x} \qquad (1)$$

We will denote equation (1) *the prior achievement interaction effect*. First, we will discuss a model suggesting how to interpret the prior achievement interaction effect, and how it compares to direct outcome measures of academic achievement. Then we will discuss the implications of unobservable heterogeneity on the estimates of the interaction effect.

A behavioural response to prior achievement

We shall start by discussing a model for the determination of educational quality as the main choice mechanism for university students. We do this by assuming that students choose programs/courses of different quality. Quality can perhaps best be interpreted as study intensity chosen by the student, and intensity will be rewarded in the labour market. However, one may also imagine that the quality dimension refers to the duration of studies, and that programmes of longer duration are higher paid. Let t denote the quality of university studies, and let achievement from university studies be given by a production function $a = a(t, x, \theta)$, where x as before denotes prior achievement, and θ is denoted study motivation. Study motivation is here used in the meaning of traits that are not directly observable to the researcher, among other things focus and persistence.[4] University achievement is assumed to

[4] See Pintrich and Schunk (2002) for a lenghty discussion on the concept of study motivation.

be increasing in all three arguments. In addition, it is natural to assume decreasing returns to quality ($a_{tt} < 0$), and it also appears logical to assume that prior achievement and study motivation boosts learning at university at the margin, i.e. $a_{tx} \geq 0$ and $a_{t\theta} \geq 0$.[5]

The cost of studying is captured by a cost function, $c(x,\theta) \cdot t$, that for simplicity is assumed to be linear in study intensity. The average cost, $c(x,\theta)$, reflects direct costs as well as psychic costs. For our purposes, we model average cost as a function of prior achievement and study motivation, assuming that the average cost is decreasing in both arguments, reflecting that the cost of study effort is decreasing in prior knowledge as well as study motivation. A university student chooses t to maximise the utility given by earnings net of the cost of studying, i.e. $u^e = y(a(t,x,\theta),x) - c(x,\theta) \cdot t$. The first-order condition for this problem is $u_t^e = y_a(\cdot)a_t(\cdot) - c(\cdot) \equiv 0$, where the first term reflects the returns to quality (the marginal returns to achievement times the marginal product of study intensity) and the second term the marginal cost of quality. To characterise how the intensity of university studies varies with prior achievement and study motivation, the first-order condition is differentiated with respect to t, x, and θ to get the comparative statics results $\partial t / \partial k = -u_{tk}^e / u_{tt}^e$, $k = x,\theta$. Note that $\operatorname{sign}(\partial t / \partial k) = \operatorname{sign}(u_{tk}^e)$ since the second-order condition for maximum is negative by assumption.

Starting with prior achievement, the derivative of interest is $u_{tx}^e = y_a^e a_{tx} - c_x + (y_{ax}^e + y_{aa}^e a_x) a_t$. The two first terms on the right hand side are positive. The first term reflects that the marginal effect of intensity on achievement is increasing in prior achievement and the second term that the cost of education is decreasing in prior achievement. The term within brackets cannot be signed a priori. This part reflects how the marginal return to achievement is affected by prior achievement. First, higher prior achievement may increase or decrease intensity depending on whether the labour market valuation of the two types of merits are regarded as complements or substitutes. One would expect this term to be positive if education at secondary school complements education at the university level. However, if employers place value on the highest-level degree only, then this term may be zero, and it may even be negative. The second term is negative and reflects the assumption that the marginal return to academic achievement is a decreasing function. The effect of study motivation is very similar to the

[5] Lowercase letters denote partial derivatives.

effect of prior achievement. The derivative is $u^e_{tx} = y^e_a a_{t\theta} - c_\theta + y^e_{aa} a_t a_\theta$, which apart from the complementary effect has similar interpretations. To sum up, prior achievement and study motivation increases study intensity since they decrease the cost of education and since they boost the marginal productivity of university quality. A counteracting factor is the direct effect on achievement making the marginal productivity of intensity less attractive because of decreasing returns to education. This means that it is very likely that prior achievement and study motivation have the same qualitative effect on study intensity. It is also likely that the intensity is increasing in x and θ, although we cannot rule out the opposite on theoretical grounds.

Denoting by $t(x,\theta)$ the choice of study intensity, earnings of an entrant can be expressed as $y^e = y\left[a\left(t(x,\theta),x,\theta\right),x\right]$, and the earnings of a non-entrant is $y^0 = y[0,x]$. Let us now turn back to equation (1). The derivative can be expressed as

$$\frac{\partial \Delta y}{\partial x} = y^e_a\left[a_t t_x + a_x\right] + \left[y^e_x - y^0_x\right] \qquad (1')$$

The effect of prior achievement on the earnings difference has two sources. The first part on the right hand side of (1') is related to the marginal valuation of the university education in the labour market. The two terms within the first brackets are the effects on academic achievement. This is what is measured when using achievement measures, such as GPA to assess study success. Prior achievement boosts achievement at the university directly (the second term within the first bracket) and indirectly via the choice of study intensity (the first term). As long as prior achievement increases study intensity then both these terms are positive. Academic achievement is then valued at the margin (captured by the term y^e_a). Note that, as long as a positive valuation is placed on academic achievement in the labour market, we expect a positive association between traditional achievement measures and earnings.

The second part on the right hand side reflects the difference in the (marginal) valuation of prior achievement in the labour market between the two groups. As before, the sign and size of this effect depends on how the marginal return with respect to prior achievement depends on academic achievement, and this in turn depends on how employers value the two forms of education; how valuable is a marginal increase in prior GPA if I have a university diploma versus how valuable is a GPA increase if I do not? From the discussion above it is clear that one should not interpret the prior achievement interaction effect as a structural parameter as it is composed of different effects that may vary over time and over different subgroups. Using earnings information and information on prior achievement, it is straightforward to obtain an estimate of the total effect, and it is also possible to obtain an estimate of the effect of prior achievement among non-entrants (y_x^0). In order to disentangle the other effects, however, one would need information that is usually not observed on a regular basis; e.g. the study effort put in by students or employer valuation of prior GPA among university entrants.

Problems with interpreting the interaction effect

One important question here is to what extent a measure of the interaction effect reflects the true dependence between prior achievement and income differences (as well as prior achievement and academic achievement). If the purpose of the analysis is to serve as a foundation for policy analysis, for example if upper secondary school grades successfully can be used to restrict access to higher education, this issue is of vital importance. Here, we will discuss two types of biases in the estimate of the interaction effect that may potentially arise. Both of them have to do with selection into higher education because of factors that researchers cannot observe. First, because of self-selection, study motivation will be unequally distributed among entrants with different prior achievement. Second, since prior achievement is correlated with study motivation at the pre-university level, an estimate of the dependence between prior achievement and the earnings difference may be biased simply because study motivation at the two educational levels are correlated.

To see the first type of bias, suppose an individual will enter university if the utility of entering is at least as large as the utility of not entering ($u^e \geq u^0$). Let the pair $\left(\hat{x},\hat{\theta}\right)$ define a marginal individual that is indifferent between entering higher education and continuing to the labour market directly after high school, i.e. $u^e = u^0$ or

$$y\left(a\left[t(\hat{x},\hat{\theta}),\hat{x},\hat{\theta}\right],\hat{x}\right)-c\left(\hat{x},\hat{\theta}\right)\cdot t\left(\hat{x},\hat{\theta}\right)-y\left(0,\hat{x}\right)\equiv 0\,.$$

Differentiating this expression with respect to x and θ and rearranging, one obtains

$$\left.\frac{\partial\theta}{\partial x}\right|_{x=\hat{x},\theta=\hat{\theta}} = -\left(\frac{y_a^e a_x - c_x \cdot t + y_x^e - y_x^0}{y_a^e a_\theta - c_\theta \cdot t}\right) \qquad (2)$$

A sufficient condition for the term within brackets in (2) to be positive is that $y_x^e - y_x^0 \geq 0$. This difference is related to the valuation of the two types of degrees as complements, i.e. if $y_{ax} > 0$. Thus, for example, if the two types of education are valued independently of each other, then $y_x^e - y_x^0 = 0$ and the condition in (2) is negative. This would mean that students with a high prior achievement that do not enter university have low study motivation and vice versa; students with low prior achievement that enter university have high study motivation. If we for the moment assume that study motivation is uniformly distributed across all prior achievement levels, then the average study motivation is falling in prior achievement among the entrants into university, meaning that study motivation and prior achievement is negatively correlated in the group of entrants. Therefore, if one fails to take study motivation into account in an empirical analysis, the estimate of the interaction effect is likely downward biased as long as study motivation has a positive impact on the interaction effect.

The second type of bias suggested arises if study motivation at the two different levels of education is correlated. Suppose that prior achievement to some extent depends upon study motivation at the pre-university level. Then if study motivation at the two levels is positively correlated as we expect, prior achievement will be positively correlated with study motivation at the academic level. Therefore, unobserved motivation will in this case give rise to an upward bias of the estimate of the interaction effect. The net effect of the two causes means that failing to take unobserved motivation into account, the interaction effect may be biased upwards as well as downward. It is very difficult to deal with the bias caused by study motivation. To some extent one may of course capture study motivation by controlling for traits known to be related to motivation, e.g. social group and parental education. However, it is likely to differ within these groups as well, implying that control variables are unlikely to eliminate the bias all together. Therefore, one should be careful in using estimates of the interaction effect for the purpose of policy analysis.

3. Data and Methodology

Selection of data

Register data from Statistics Sweden is used to study the earnings effects from entrance into universities. The sample drawn consists of individuals born between 1972 and 1974 who graduated from academic track programs in Swedish upper secondary schools. In Sweden, a large majority of teenagers enter upper secondary school at the age of 16. Among the individuals born 1972 to 1974, nearly 80 percent had graduated from upper secondary school by 1996. Out of these, 46 percent (127 000 individuals) had attended a program on academic track. The outcome variable used is the sum of income from employment and self-employment in the year 2002. This means that our study measures the incomes of individuals as they are in the age span 28 to 30 years. This is a reasonable measure of early career outcomes for individuals that attained higher education considering that the large majority of the students that go on to university do so either directly after finishing upper secondary schools or in their early twenties.

Our data regarding the registration of individuals into the universities cover the years up to 2002. In order to make a fair comparison between university entrants and non-entrants, only university entrants that were registered at a Swedish university before the year 1996 are included. Students that were registered first time between 1996 and 2001 are excluded. There are two main reasons for this delimitation. First, one must allow a sufficient amount of time to pass between university entrance and observation in the labour market, so that the individual has the opportunity to finish higher education. Second, since we are interested in earnings differences in the different parts of the grade distribution, we would not like the estimates to be influenced by the possibility that individuals with low grades enter universities at a later date than high-performing individuals. Indeed, if we look at the data, the university entrance date is negatively correlated with the GPA in upper secondary school. Including only those university students that entered before 1996 means that the students have had at least seven years from their university registration date to the year when we observe their income.

The comparison group (non-entrants) is formed by including the individuals that did not register into a university before the year 2002. Moreover we require the individuals to be registered as employed in the autumn of 2002. All together this means that out of the 127 000 students that graduated from upper secondary school academic track programs, 15 000 students are dropped as they registered into a university after 1995, and approximately 19 000 are dropped since they were not working in the autumn of 2002, leaving 93 268 observations in the final sample. The number of female students is slightly larger (48 107) than the number of male students (45 161). In the estimations, observations with missing information are excluded, meaning that the sample sizes are substantially lower when we control for background information. Definitions of the variables used in the empirical analysis are presented in Table A1 in the Appendix, and descriptive statistics is presented in Table A2.

GPA quintile groups

As mentioned in the preceding section, the main research strategy is to compare the income differences between university entrants and non-entrants over the upper secondary grade point averages, and we do this by comparing students in different GPA-groups. This section contains a description of different GPA groups where students have been divided into five groups of equal size, where the 20 percent lowest achieving form group one. Male students and female students are treated separately. Forming quintile groups using grade point averages will create a large number of ties. Rather than splitting ties to create groups of exactly equal size, we choose the convention to include the upper limit value in the lower group. In particular, this means that the lowest achieving group will exceed 20 percent and the highest achieving group will contain less than 20 percent of the observations. However, in no case the smallest group will contain less than 18 percent of the observations.

In the norm-referenced grading system that these students were graded according to, GPA varies between 1 and 5. Table 1 shows the percentage of university entrants in the different quintile groups, as well as the relative size and the upper limit associated with each group in our sample of data.

Table 1 HERE

As can be seen from the table, very few of the top-performing students in academic track programmes choose not to enrol in university studies in Sweden. Among the top-performing men 98 percent are university entrants, and 97.5 percent of the females are registered at a university. Among the low-achievers, the proportions of university entrants are lower; approximately 30 percent of the men and 36 percent of the females.

Table 2 HERE

2002 incomes among the university entrants and non-entrants are shown in Table 2. The average female income in the sample was 213 000 SEK in 2002. One can note that the average income of high-performing females that were not enrolled in a university did not reach the average income level, meaning that most females that did not enter into a university had below the average income. Earnings tend to increase with GPA in upper secondary school. This is especially true for university-enrolled females, while the income-GPA association is much weaker for non-entrants. This suggests that the income differences between entrants and non-entrants tend to increase with the GPA in upper secondary school.

The male income also increases with the GPA in upper secondary school. However, in contrast to the incomes among women, the highest performing non-entrants (quintile 5) earn well in excess of a large majority of the university-enrolled men. Another observation worth mentioning is that the relative income differences between entrants and non-entrants are smaller than the corresponding figures for the women. Finally one should note that there are large income differences between males and females at this age.

Estimation strategy

The effects of education on subsequent incomes are usually studied by estimating the returns to education (Mincer, 1974). A basic idea is that education is an investment that will yield pay-off in the form of improved future labour market outcomes. The by now standard method to address the returns to education is to compare the incomes of individuals with different years of completed education or those with and without a specified education. However, since our main purpose is to discuss the differences with respect to university entrance, the concept of returns to education in this case means the returns to university entrance, thus not conditioning on programme completion. Let y denote income for individual i. The baseline empirical specification is

$$\ln y_i = f\left(ENTR_i, GPA_i\right) + \delta GPA_i + z_i + \varepsilon_i \tag{3}$$

where the dependent variable is the logarithm of income. In all specification, we assume that income develops linearly with the upper secondary school GPA, $f\left(ENTR_i, GPA_i\right)$ is some function (to be discussed below) of university entrance and the grade point average from upper secondary school capturing the prior achievement interaction effect, z denotes a set of background variables and associated parameters capturing the influence of individual traits and differences in upper secondary schooling. Finally, ε is an error term capturing unobserved heterogeneity. Of key interest here is the interaction between the upper secondary school GPA and entrance on income. Since the function f is unknown a priori, it is important to estimate its' parameters using a flexible functional form. We therefore choose to work with three different functional forms. As a first specification, the samples of individuals are divided into different GPA groups, and between group differences are estimated.

In a second specification, f is modelled as a piecewise linear function (linear spline function) for the income differences, and in the third type of specification, the interaction effect potentially develops non-linearly using higher order polynomials. See equation A1-A3 in the Appendix for details. As a general estimation strategy, for each of the specifications, we start by estimating a general model and then test if restricting some or all of its parameters can reduce the model.

Covariates

A number of covariates are used in order to take heterogeneity, such as unobserved motivation differences, into account, as well as controlling for income differences not necessarily related to previous achievement. Apart from the upper secondary school GPA, we impose controls for age, immigration status, parental education, and primary school grades. At the time the individuals in the sample studied at upper secondary school, students on academic track could participate in five different programmes. Using dummy variables, we control for the upper secondary school programme attained. The major reason for including controls for educational orientation is that grading in upper secondary school may differ between upper secondary school programmes.

4. Results

GPA quintile specific income differences

Let us start by presenting the results from estimating the regression model using the step-wise interaction model presented in equation 4. In order to obtain a suitable empirical specification, the male and female samples are first divided into deciles with respect to upper secondary GPA, and equations with ten level effects and interaction effects are estimated. We then test if the model can be reduced to five quintile groups with separate interaction effects. It turns out that five interaction effects describes the data well, so in the continuation quintile specific interaction effects are presented. In Table 3, the results from the female income equations are presented, and the corresponding estimates for the males are presented in Table 4. The first column shows a parsimonious version of the model, where the only control variable is the upper secondary GPA, and the quintile group level effects. In column II background control variables are added, and in the third column upper secondary school educational orientation is added as well. To save on space, only the estimated income differences by GPA quintile group are presented. A full set of results can be obtained from the authors.

Table 3 HERE

Table 4 HERE

Introducing background and educational orientation controls have only small effects on the estimated income differences, indicating that background controls do not appear to be very important for the interaction between entrance and GPA. In the male models (Table 4), the estimates are lowered somewhat for the two top quintile groups when introducing background controls, indicating that unobserved motivation may render the interaction effects to biased downwards, while in the female models, if anything, the difference is the opposite. In the female model specifications, the income differences between entrants and non-entrants are found to be positive and statistically significant for all quintile groups. Entrants in the lowest achieving group earn approximately six percent more than non-entrants, and this difference increases with prior achievement for each of the subgroups up to 32 percent in the top quintile group. These results indicate that there may be large differences in the returns to entrance among the females on academic track. The male earnings differences (see Table 4) are much smaller than the female differences. The lowest 60 percent of the males earns approximately the same independently of whether they entered a university or not. The difference is approximately eight percent in the fourth quintile and ten percent in the top quintile; however the latter is not statistically significant.

A common feature of the males and female samples studied here is that the income differences between entrants and non-entrants appear to grow with the upper secondary GPA, indicating that the income-GPA profiles are steeper among the entrants. In order to test formally if there are group-wise differences, and if income differences are growing in GPA, we proceed by testing hypotheses of no differences between various subgroups using Wald tests. The results are presented in Table 5 and based on the estimates of column III in tables 3 and 4. The first row of Table 5 tests the null hypothesis of no differences between the two bottom quintiles, the second row between the three bottom quintiles, and so on. Starting with the female income differences, we find that equality of the income differences in the two bottom quintiles are clearly not rejected, while equality among the three bottom quintiles is marginally accepted using the 95-percent level as the cut-off. However, we can reject equality among the four lowest or all five of the quintiles. Together with the fact that the estimated interaction effects are growing in GPA, these results imply that the income difference grows with the GPA.

Table 5 HERE

Applying the same tests on the male estimates give results that are similar to the female income differences. Here, equal income differences are rejected among the lowest four or all five of the quintile groups, and we may interpret the association between income differences and GPA as an increasing relation. One interesting aspect of these results is that, the income differences appear to be approximately constant among the lower achieving groups. In particular, we do not find big differences between the lowest achieving and those in the middle of the grade scale. The lowest achieving group are those with particularly poor grades and one would expect them to have worse outcomes in terms of the return to entrance. One should also note that in the two top quintiles, more than 90 percent entered a university during the period studied. Thus, the private value of entering a university is well captured by this group.

Alternative specifications

Estimating between group differences may not capture all the aspect of the distribution of entrance returns with respect to prior achievement. Let us therefore present the estimates of the alternative specifications. The numerical results are presented in the appendix. First, the piecewise linear interaction model, defined in equation (A2), is estimated by dividing the male and female samples into deciles by GPA and then estimating slope parameters for each of the ten groups. A general to specific testing procedure is used in this case, where the parameter with the lowest level of significance is restricted to zero, and then the model is re-estimated. The procedure is repeated until all the remaining parameters are significantly determined at the 95 percent level. In a similar way the polynomial interaction model (see equation A3) is estimated. Here, we start by assuming a polynomial expression of order nine, and then reduce the model by restricting the highest order parameter to zero until the remaining parameters are significantly determined. The results from the piecewise linear interaction model are presented in Table A3, and the results from the polynomial model in Table A4. Since both the piecewise linear interaction model and the polynomial interaction model contain the linear interaction model as a special case, we also display the results from the latter type of model in Table A4. All the estimated models contain controls for background variables and economic orientation. A diagrammatic exposition of the results these three models are given in Figures 2 and 3.

Figure 2 HERE

Figure 3 HERE

Estimated female income differences over the GPA distribution are displayed in Figure 2. In the linear interaction model, the slope parameter is approximately 0.03, meaning that there is a 3 percent gain in the returns to entrance per decile increase in the GPA. The other two models predict approximately the same outcome around the middle part of the GPA distribution. However, at the endpoints, the regression line of the polynomial interaction model differs from the others. We interpret this to mean that the polynomial model does not predict the data well at the end points of the sample. Male differences are smaller (see Figure 3), the linear interaction model has a slope of approximately 0,1, and the slopes of the other two models are even flatter for the main part of the GPA distribution. Another important observation is that the regression lines in the flexible models are not monotonically increasing. However, the lowest achieving decile appear to fare worse than deciles 2 and 3, which may indicate that the lowest achievers have problems in achieving a good educational outcome.

Sample restrictions

The income of the individuals in the sample is measured at a young age. At this age many young adults form families, and one may suspect that part of the sample is not working full time for reasons related to preferences. Since one potential interest lies in answering the question if higher education admission is suitable to everyone, such voluntary restrictions to income may wrongly be interpreted as an effect of poor learning outcomes. One way to study potential problems associated with tastes for work is to reduce the sample by eliminating individuals with low income. Of course, it is difficult to choose a particular limit below which to consider individuals as voluntarily abstaining from work. We base our restriction on a comparison of earnings data with hours of work by Antelius and Björklund (2000), to choose 120 000 SEK as the limit. Performing the same type of analyses as above, the results suggest that male differences are largely unaffected, while female earnings differences as well as the slope of the regression lines become smaller. The linear interaction effect is estimated to be about half that of the estimate for the original sample. However, the shape of the interaction effect, when estimated using the non-linear specifications, is similar to the original sample, being flatter in the middle parts of the distribution and steeper at the tails.

5. Concluding remarks

This paper has presented an alternate approach to how to validate selection instruments to higher education. The idea was to avoid common problems in validation research – such as limitation of range and problematic criteria for academic success - by including also students who did not enter higher education (by rejection or own choice) and use earnings as criterion for academic success. From a theoretical point of view, there is a close connection between the traditional definition of predictive validity, measured as the correlation between university achievement and prior grades, and the association between the earnings of university entrants and prior grades. We have used labour market information to assess if prior grades predict income premiums of entering university. The main results are the following: (i) the income premium at labour market entrance for university students is positive. However, it is larger for females than for males indicating that university education is economically more important for females; (ii) the gradient of the achievement interaction effect is small for a large part of the prior grade distribution, meaning that using upper secondary GPAs to predict who will benefit the most from university entrance is not very informative for a majority of the student population; (iii) in the tails of the prior grade distribution, the gradient is larger, indicating that the lowest achieving students (in terms of prior grades) are not successful students while high-achievers are.

We also discussed possible caveats in interpreting the prior achievement interaction effect, and pointed out two possible selection mechanisms that may bias the estimated parameters, one being that there is a negative correlation, on the margin, between prior achievement and study motivation, and the other on that study motivation at different levels of the education system may be positively correlated. Together these two sources mean that one cannot generally tell if estimates of the prior achievement interaction effect are biased upwards or downward. Controlling for observable characteristics usually assumed to be correlated with motivation did not affect the results in any important way.

The results of the paper suggest some avenues for further research. First, the empirical connection between achievement at university (university GPA) and earnings is not fully understood. Where available, one should try to associate earnings with achievement at university directly to address this connection. Second, the fact that the gradient of the prior achievement interaction effect is large at the lower tail of the prior grade distribution makes it interesting to study more thoroughly the academic success of poorly performing students. At the least, this is interesting from a policy point of view, since cuts in the number of student positions will affect this group the most. Third, our results are not directly applicable in a cost-benefit analysis of higher education; higher performing students generally attain programs that involve higher direct costs as well as longer duration. In addition, wages at labour market entrance do not capture the entire effect from higher education over the entire life span. Therefore, to get a better estimate of the surplus from higher education, the costs and the program duration, as well as the income path, should be taken into account. Finally, there is an interest in comparing entrants with non-entrants in particular programs to see if different programs differ in the achievement interaction effect. In this paper, we did not have information about applicants. Future studies may well have access to such information.

Appendix

Specifications of the prior achievement interaction effect $f(\cdot)$

(i) Group differences

$$f\left(ENTR_i, GPA_i\right) = \sum_{j=1}^{J}\left(\alpha_j ENTR_i + \rho_j\right) \quad \text{(A1)}$$

In this specification, the interaction effect is modelled in a stepwise fashion. Here, j denotes the J different GPA groups, ρ_j is a between group deviation from the linear association between income and GPA, and α_j the percentage income difference between university entrants and non-entrants in group j.

(ii) Piecewise linear function

A linear spline function for the income differences is given by

$$f\left(ENTR_i, GPA_i\right) = \left(\beta_0 + \beta_1 GPA_i + \sum_{j=2}^{J}\beta_j\left(GPA_i - \tilde{G}_{j-1}\right)D_j\right)ENTR_i \quad \text{(A2)}$$

where J is the number of segments of the piecewise linear function, D_j is a dummy variable taking the value one if $GPA_i - \tilde{G}_{j-1} > 0$ and zero otherwise, and $\tilde{G}_{j-1}$ is the (upper) threshold value of segment *j-1*. This means that for the lowest GPA group the estimated income difference between entrants and non-entrants is $\beta_0 + \beta_1 GPA_i$, for the second lowest group $\beta_0 + \beta_1 GPA_i + \beta_2\left(GPA_i - \tilde{G}_1\right)$, and so on.

(iii) Polynomial function

In this case, the function *f* is specified as

$$f\left(ENTR_i, GPA_i\right) = \left(\gamma_0 + \sum_{k=1}^{K}\gamma_k GPA_i^k\right)ENTR_i \quad \text{(A3)}$$

where the expression within parenthesis captures the interaction effects, and where K denotes the order of the polynomial.

Tables

Table A1: Variable definitions

Variable	Definition
Logarithm of income	The natural logarithm of income from employment and self-employment in 2002
Age	The age of the individual defined as 2003 minus the year of birth
Father upper secondary	Dummy variable taking the value one if the father's highest level of education is upper secondary education
Father post secondary	Dummy variable taking the value one if the father has some postsecondary education and zero otherwise
Mother upper secondary	Dummy variable taking the value on if the mother's highest level of education is upper secondary education
Mother post secondary	Dummy variable taking the value one if the mother has entered postsecondary education and zero otherwise
First generation immigrant	Dummy variable taking the value on if the individual and both parents are born outside of Sweden
Second generation immigrant	Taking the value on if the individual is born in Sweden and both parents are born outside of Sweden
Upper secondary GPA	The grade point average from upper secondary education. GPA is measured on a scale one to five, where one is the lowest grade and five is the highest grade. In the empirical analysis students are percentile ranked according to GPA and placed on a scale zero to ten.
Secondary GPA	The grade point average from upper secondary education. Grades are given on a scale one to five, where one is the lowest grade and five is the highest grade.
Educational orientation	Upper secondary school programme. Educational orientation is constructed as 5 different dummy variables taking the value one if the individual attained a particular programme, and where the Social science programme is the reference group.

Table A2: Descriptive statistics for the key variables

Variable	Females		Males	
	Mean	N	Mean	N
Income	212.8	48 107	299.5	45 161
	(110.4)		(144.2)	
Age	29.00	48 107	29.02	45 161
	(0.82)		(0.82)	
Father upper secondary	0.308	39 414	0.308	37 679
	(-)		(-)	
Father post secondary	0.361	39 414	0.402	37 679
	(-)		(-)	
Mother upper secondary	0.295	44 242	0.280	41 312
	(-)		(-)	
Mother post secondary	0.382	44 242	0.421	41 312
	(-)		(-)	
First generation immigrant	0.023	47 575	0.025	44 629
	(-)		(-)	
Second generation immigrant	0.029	47 575	0.027	44 629
	(-)		(-)	
Upper secondary GPA	3.515	48 107	3.387	45 161
	(0.616)		(0.658)	
Upper sec GPA, rank	5	48 107	5	45 161
Primary school GPA	3.914	47 128	3.757	44 324
	(0.787)		(0.775)	

Note: Standard deviations are given within parentheses

Table A3: Income differences, piecewise linear interaction effects

Parameter	Females		Males	
	Full model	Restricted model	Full model	Restricted model
β_0	-0.061	-	-0.119*	-0.088**
	(0.071)		(0.051)	(0.028)
β_1	0.172*	0.103**	0.123*	0.075**
	(0.087)	(0.026)	(0.062)	(0.017)
β_2	-0.208	-0.128**	-0.075	-
	(0.201)	(0.042)	(0.083)	
β_3	0.013	-	-0.072	-0.118**
	(0.077)		(0.046)	(0.030)
β_4	0.154*	0.153**	0.018	0.050**
	(0.066)	(0.045)	(0.041)	(0.017)
β_5	-0.191**	-0.181**	0.016	-
	(0.059)	(0.050)	(0.036)	
β_6	0.120*	0.087**	-0.009	-
	(0.053)	(0.027)	(0.003)	
β_7	-0.063	-	-0.001	-
	(0.048)		(0.003)	
β_8	0.053	-	0.016	-
	(0.048)		(0.032)	
β_9	-0.004	-	-0.002	-
	(0.048)		(0.031)	
β_{10}	0.088	0.113**	0.081*	0.102**
	(0.049)	(0.027)	(0.037)	(0.020)
R^2	0.053	0.053	0.067	0.067

Note: The parameter estimates are based on the piecewise linear interaction model presented in equation A2. β_0 denotes the constant interaction parameter and β_1-β_{10} the slope changes in comparison to the preceding decile group. Standard errors are presented within parenthesis. ** denotes statistically significant parameters on the 99 percent level, * significance on the 95 percent level. The number of observations is 37 135 and 35 181 in the female and male equations, respectively.

Table A4: Income differences, polynomial interaction effects

Parameter	Females		Males	
	7th grade	Linear model	5th grade	Linear model
γ_0	-0.170 (0.096)	0.005 (0.019)	-0.151** (0.045)	-0.012 (0.011)
γ_1	0.723** (0.265)	0.030** (0.005)	0.252* (0.068)	0.010** (0.003)
γ_2	-0.710** (0.261)	-	-0.120** (0.036)	-
γ_3	0.323** (0.120)	-	0.026** (0.008)	-
γ_4	-0.076** (0.029)	-	-0.003** (0.001)	-
γ_5	0.010** (0.003)	-	$9.67 \cdot 10^{-5}$** ($3.17 \cdot 10^{-5}$)	-
γ_6	$-6.36 \cdot 10^{-4}$** ($2.64 \cdot 10^{-4}$)	-	-	-
γ_7	$1.68 \cdot 10^{-5}$** ($7.20 \cdot 10^{-6}$)	-	-	-
R^2	0.053	0.052	0.067	0.066

Note: The parameter estimates are based on the polynomial interaction model presented in equation A3. γ_0 denotes the constant interaction parameter and γ_1-γ_{10} linear and higher order interaction parameters. Standard errors are presented within parenthesis. ** denotes statistically significant parameters on the 99 percent level, * significance on the 95 percent level. The number of observations is 37 135 and 35 181 in the female and male equations, respectively.

References

Antelius, J., and Björklund, A. (2000). How Reliable are Register Data for the Return on Schooling? An Examination of Swedish Data. *The Scandinavian Journal of Educational Research, 44*(4), 341-355

Cohn, E., Cohn, S., Balch, D.C., and Bradley Jr, J. (2004). Determinants of undergraduate GPAs: SAT scores, high-school GPA and high-school rank. *Economics of Education Review 23*, 577-586

Elias, P., and Purcell, K. (2004). Is Mass Higher Education Working? Evidence from the Labour Market Experience of Recent Graduates. *National Institute Economic Review 190*, 59-74

Gregory, R.J. (2004). *Psychological testing: history, principles and applications* (4th edition) . Needham heights: Allyn & Bacon.

Henriksson, W., and Wolming, S. (1998). Academic performance in four study programmes: a comparison of students admitted on the basis of GPA and SweSAT-scores with and without credits for work experience. *Scandinavian Journal of Educational Research, 42*(2), 135-150

Lyrén, P-E. (2008). Prediction of Academic Performance by Means of the Swedish Scholastic Assessment Test, *Scandinavian Journal of Educational Research, 52*(6), 565-581

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational Measurement* (3rd edition). New York: Macmillan.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: NBER Press.

Osborne, M. (2003). Increasing or Widening Participation in Higher Education? – a European overview, *European Journal of Education, 38*(1), 5-24

Psacharopoulos, G., and Patrinos, H.A. (2004). Returns to Investment in Education: A Further Update, *Education Economics, 12*, 111–134

Pintrich, P.R., and Schunk, P.D. (2002). *Motivation in education: theory, research and applications* (2nd ed.). Upper Sadlle River, NJ: Pearson Education

Svensson, A., Gustafsson, J-E., and Reuterberg, S-E. (2001). *Högskoleprovets prognosvärde.* [The predictive value of the SweSAT] Högskoleverkets rapportserie 2001:19R. Stockholm: National Agency for Higher Education.

Svensson, A., and Nielsen, B. (2004). *Urvalet till juristutbildningen. Skillnader mellan dem som antas i betygs- respektive provurvalet.* [Selection to law school. Differences between those admitted on the basis of grades and test] (IPD-rapport nr 2004:11). Göteborg: Gothenburg University, Institutionen för pedagogik och didaktik.

Walker, I., and Zhu, Y. (2008). The College Wage Premium and the Expansion of Higher Education in the UK. *Scandinavian Journal of Economics 110*(4), 695-709

Wolming, S. (1999). Validity issues in higher education selection: a Swedish example. *Studies in Educational Evaluation, 25*(4), 335-351

Wolming, S. (2001). Att värdera urvalsinstrument. Några reflektioner över begränsningar och möjligheter [Validating selection instruments. Some reflections over limitations and possibilities]. *Pedagogisk forskning i Sverige, 6*(2), 122–130

Wolming, S. & Wikström, C. (2010). The concept of validity in theory and practice. *Assessment in Education – principles, policy and practice, 17*(2), 117-132

Table 1: The percentage of university entrants in GPA quintile groups

GPA quintile group	Entrants	Group size	Limit GPA
Females			
Quintile 1	36.0	22.4	3.00
Quintile 2	65.9	18.2	3.33
Quintile 3	79.5	19.6	3.67
Quintile 4	90.1	21.3	4.07
Quintile 5	97.5	18.5	5.00
All	73.0	100.0	
Males			
Quintile 1	30.8	20.8	2.82
Quintile 2	59.9	19.7	3.19
Quintile 3	79.0	20.4	3.53
Quintile 4	91.4	20.9	4.00
Quintile 5	97.9	18.2	5.00
All	71.2	100.0	

Table 2: Income in GPA quintile groups, 1000SEK

GPA quintile group	Entrants	Non-entrants
Females		
Quintile 1	187.5	176.0
	(93.4)	(92.6)
Quintile 2	197.7	184.7
	(97.3)	(97.6)
Quintile 3	210.9	195.5
	(101.9)	(103.9)
Quintile 4	229.7	199.7
	(111.2)	(115.1)
Quintile 5	262.1	199.2
	(127.7)	(112.0)
Males		
Quintile 1	271.3	260.5
	(114.2)	(103.0)
Quintile 2	284.3	281.6
	(112.2)	(128.9)
Quintile 3	298.2	289.0
	(143.9)	(112.8)
Quintile 4	313.8	293.7
	(185.2)	(125.0)
Quintile 5	347.2	332.8
	(200.0)	(143.5)

Note: Average income from employment and self-employment. Standard deviations within parentheses.

Table 3: Income Differences (per cent) over Quintile GPA Groups; Females

Estimate	I	II	III
Quintile 1 (α_1)	0.069**	0.063**	0.057**
	(0.018)	(0.021)	(0.021)
Quintile 2 (α_2)	0.093**	0.098**	0.099**
	(0.020)	(0.024)	(0.024)
Quintile 3 (α_3)	0.124**	0.126**	0.135**
	(0.022)	(0.026)	(0.026)
Quintile 4 (α_4)	0.201**	0.189**	0.202**
	(0.031)	(0.034)	(0.034)
Quintile 5 (α_5)	0.330**	0.312**	0.324**
	(0.064)	(0.072)	(0.072)
Background controls	No	Yes	Yes
Educ orientation	No	No	Yes
R2	0.040	0.044	0.053
NOBS	48105	37135	37135

Note: The dependent variable is the logarithm of income from employment and self-employment. Robust standard errors (White) are presented within parentheses.

Table 4: Income Differences (per cent) over Quintile GPA Groups; Males

Estimate	I	II	III
Quintile 1 (α_1)	0.017	0.008	0.0002
	(0.011)	(0.013)	(0.013)
Quintile 2 (α_2)	0.011	0.028*	0.017
	(0.011)	(0.012)	(0.012)
Quintile 3 (α_3)	0.018	0.032*	0.023
	(0.013)	(0.013)	(0.013)
Quintile 4 (α_4)	0.064**	0.086**	0.082**
	(0.019)	(0.021)	(0.022)
Quintile 5 (α_5)	0.038	0.089	0.102
	(0.052)	(0.060)	(0.060)
Background controls	No	Yes	Yes
Educ orientation	No	No	Yes

R2	0.032	0.040	0.068
NOBS	45160	35181	35181

Note: The dependent variable is the logarithm of income from employment and self-employment. Robust standard errors (White) are presented within parentheses.

Table 5: Wald Tests of Restrictions on Income Differences over GPA Quintile Groups

Null hypothesis	Wald statistic Females	Males
$\alpha_1=\alpha_2$	1.78 (0.182)	0.93 (0.335)
$\alpha_1=\alpha_2=\alpha_3$	2.84 (0.059)	0.86 (0.422)
$\alpha_1=\alpha_2=\alpha_3=\alpha_4$	5.06** (0.002)	3.69* (0.011)
$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5$	6.01** (0.001)	3.23* (0.012)

Note: The tests are based on the estimates of specification III in Tables 3 and 4. The test is distributed as F(r,n-q) where r is the number of restrictions and n-q the number of degrees of freedom. P-values for accepting the null are presented within parenthesis. ** the null is rejected at the 99-percent level, * the null is rejected at the 95-percent level.

Figure 1: The relationship between selection instrument and criteria for academic success, according to Gregory (2004)

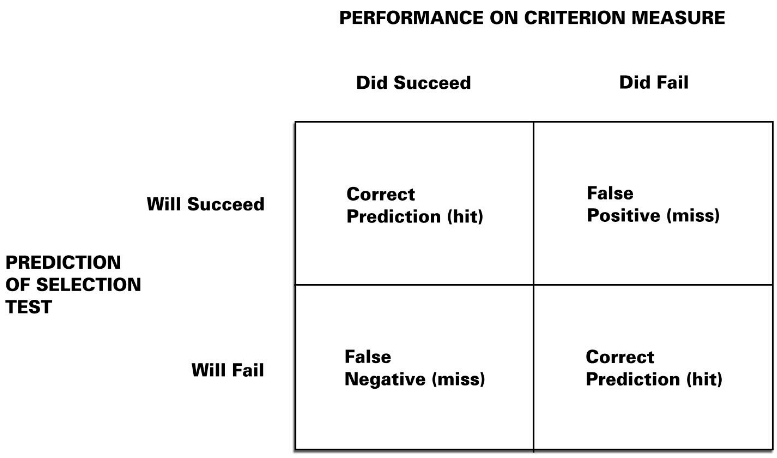

Figure 2: Estimated income differences; females

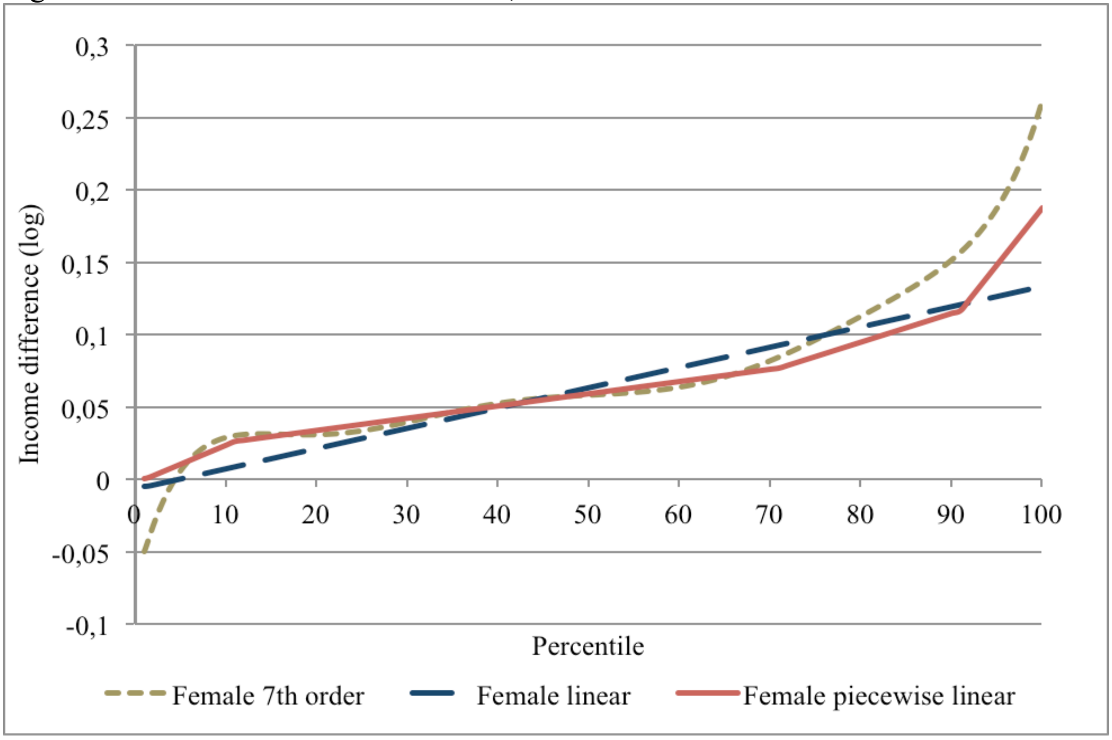

Figure 3: Estimated income differences; males