



IZA DP No. 5714

**Do They Understand the Benefits from Education?
Evidence on Dutch High School Students'
Earnings Expectations**

Jacopo Mazza
Joop Hartog

May 2011

Do They Understand the Benefits from Education? Evidence on Dutch High School Students' Earnings Expectations

Jacopo Mazza
University of Amsterdam

Joop Hartog
*University of Amsterdam
and IZA*

Discussion Paper No. 5714
May 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Do They Understand the Benefits from Education? Evidence on Dutch High School Students' Earnings Expectations*

Using an internet collected dataset, we will provide some empirical evidence on the information that Dutch high school students possess before their decision on tertiary education participation. The sample is prone to selective participation and high attrition, but we detect little systematic effects and inconsistent reporting of probability distribution is not more frequent than in controlled settings. We find little support for patterns that one would expect from individuals having private information. Girls expect substantially lower earnings in all schooling scenarios, but implicit rate of returns do not differ from those anticipated by boys. We find no evidence of expected risk compensation.

JEL Classification: J24, J31, I20

Keywords: wage expectations, wage risk, risk compensation

Corresponding author:

Jacopo Mazza
Department of Economics
University of Amsterdam
Roeterstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: j.mazza@uva.nl

* For computational help with Matlab, the authors thank Domenico Massaro.

1. Introduction

Human capital theory defines education as an investment in someone's own capacity, his human capital exactly. This theory supposes that individuals choose rationally the most attractive alternative among different educational levels and types, given their needs and preferences. Rational investment postulates informed agents and investment in human capital makes no exception. Neither should the theory of human capital ignore risk, a feature so prominent in investment theory.

Unfortunately little is known in the literature on the information quality of this particular category of investors – students – at the time of choice. Hereafter, we will provide some empirical evidence on the level and extent of knowledge that Dutch high school students possess at the time immediately preceding the most relevant education decision: proceeding or not to tertiary education. We will also assess if they are aware of the risky nature of such an investment and if, and to which amount, they require a compensation for it. We will further extend our analysis to check whether students possess private information that they exploit when forming their expectations and which factors influence their probabilistic expectations. The particular structure of our data will allow us to shed some light on a longstanding question in labor economics: calculating the expected returns on education.

We used a freely accessible internet survey to collect data on individuals' expected median wages for different schooling scenarios and we apply a method developed by Dominitz and Manski (1996) to elicit the individuals' perceived dispersion of earnings. So far, this method has only been implemented in a controlled environment, under supervision and with direct feedback on inconsistent probability statements¹. To our knowledge this is the first study concerning students' wage expectations by an open inquiry at the

¹ The Dominitz and Manski method was copied by Wolter (2000) and by Schweri et al (2008).

internet, covering many different studies, rather than a supervised data collection for a narrowly defined population. We will pay due attention to the pros and cons of this specific method, before analyzing the observations.

As mentioned, the literature directly eliciting students' expectations is quite scarce. Blau and Farber (1991) asked a sample of American college students to give a direct forecast of their entrance wage and their wage after ten and twenty years. Betts (1996) exploits another sample of undergraduate American college students to examine their knowledge about current salaries by type of education and to link the level of knowledge to specific individual characteristics concluding that the information possessed by students is rather accurate and specialized for field of interests. More recently, Zafar (2011) collected college students expectations about their major specific outcomes finding that even if prone to overconfidence about their academic performance, students update their expectations in expected ways. We follow Dominitz and Manski (1996), who paid special attention to anticipated earnings dispersion. They elicited, from a small sample of both high school and college students, their income expectations at different moments of their life cycle conditional on different schooling scenarios. They also extrapolate a subjective earnings distribution, characterizing in this way the riskiness of educational investment. They find positive perceived returns to college education and awareness of the uncertain nature of the investment in their particular sample of students. Their main contribution resides in the methodological part of data collection. They clear some widespread skepticism on the use of expectation data. As they note, this antipathy, even though longstanding, is based on a narrow foundation². In fact, they show that in their sample the internal consistency of the answers and of the variation of responses is rather high.

Our use of high school students has the considerable advantage of targeting on the most relevant group of people, at the decisive moment for their educational

² For a history on probabilistic expectations elicitation and their use in modern economic theory see, Manski (2004).

investment decision. Undeniably, the decision on undertaking tertiary education is the most relevant one in the process of human capital accumulation. Therefore, shedding some light on the information set and awareness of this particular group is of great interest in itself. The design of our survey will also allow us to check if the level of information increases as the time of choice approaches, as the human capital theory would suggest. An additional contribution of our work will be that of showing some findings about the role that perceived ethnic discrimination on the labor market might play already in the formation of students' expectations.

2. The Dutch educational system

The Dutch schooling system³ is organized in three cycles: elementary school (*Basisonderwijs*), high school (*Voortgezet Onderwijs*) and tertiary education (*Vervolgonderwijs*).

Elementary school is common to all pupils and caters for children from five (sometimes four) to twelve years old. At the end of this course of study a test – the so called CITO – is administered. The score of this test, jointly with the advice of elementary school teachers, determines in which of the three types of high schools a student is allowed to enroll.

The three types of high schools provided for in this scheme are: VMBO (*voorbereidend middelbaar beroepsonderwijs*, literally, “preparatory middle-level vocational education”), HAVO (*hoger algemeen voortgezet onderwijs*, literally, “higher general continued education”) and VWO (*voorbereidend wetenschappelijk onderwijs*, literally, “preparatory scientific education”).

VMBO lasts four years, from age twelve to sixteen. The type of curricula provided combine vocational training with theoretical education in languages, math, history, arts and sciences. A VMBO diploma gives direct access only to MBO (*middelbaar beroepsonderwijs*, literally, “middle-level vocational

³ See appendix for a summarizing scheme.

education"). This type of high school is the most practical of the three. To have access either to HBO or university, the two types of institutions in tertiary education, one needs to complete his education with an HAVO or VWO diploma respectively.

HAVO takes one year longer than VMBO: from age twelve to seventeen. The five years of study are divided in two cycles. In the first three years subjects – languages, mathematics, history, arts and sciences – are common to all students and form the *Basisvorming* (literally, "basis education"). The last two years, the *Tweede fase*, are differentiated in four streams of study depending on the core subjects of a student's elected curriculum. A HAVO diploma allows direct access to HBO and from there, after a year of study, to university education.

VWO has six grades and is attended from age twelve to eighteen. VWO is divided in *Atheneum* and *Gymnasium*. The first one does not include classical languages, while the second one does. VWO guarantees direct access to university and it is mainly considered as a preparation to it.

MBO cannot be regarded as tertiary education. It is vocationally oriented and mainly leads to middle management positions either in industry or in the service or government sectors. It lasts a maximum of four years.

The proper tertiary education system is constituted by HBO (*Hoger Beroeps Onderwijs*, literally "higher professional education") and WO (*wetenschappelijk onderwijs*, literally "scientific education" or university as commonly intended). Access to HBO is possible with an MBO, HAVO and VWO diploma. The course of study takes four years (three if coming from VWO) and it awards a *bachelor* degree. An additional one or two year of study lead to a *master* degree.

University is directly accessible only from VWO or with a *propedeuse* year in HBO. WO is oriented towards theoretical higher learning in arts or sciences. It is now divided in two cycles: the first three years leading to a *bachelor* degree and an additional year (two in case of practical or research master) for a *master* degree.

3. The data

3.1 Survey method

Our survey was administered via internet from a freely accessible web page. The participation in the survey was completely voluntary. The survey was organized in collaboration with NIBUD, an institute for informing households on sound financial practice. NIBUD periodically runs a questionnaire among high school students (approximately 12-18 years old) in the Netherlands. The main objective is to learn about income and consumption behavior of the young. Next to many questions about the student's jobs and their expenditures, the students are asked to provide some background information regarding their social-cultural background (for example, age, job status of their parents, and place where they live). In the 2006 wave, all students were asked the usual NIBUD questions. After these questions, the students of the *bovenbouw* (last years of high school) *only* were asked additional questions, including those analyzed here. The link to the questionnaire was set at the NIBUD web page and at the MSN site (*Windows Messenger*⁴). Given the widespread use of this program among Dutch high school students, this gave us potential access to a wide and non selected audience. The questionnaire was made available for compilation between March and July 2006. On the internet page preceding the questionnaire it was explicitly stated that the questions were addressed specifically to high school students from whichever type of school, in the final two years.

Even though the announcement of the site assures us of access to a non-selective potential sample, our methodology might give rise to several concerns on selectivity of people effectively participating in the survey. We will try to address these concerns further on. On the other hand, a completely voluntary participation might reassure us on the carefulness of responses since it is hard

⁴ The program is freely downloadable from internet. In its *Web Messenger* version it doesn't need an installation on the computer in use, making it, thus, available also from shared or public computers such as those available in public libraries or high schools computers.

to believe that a teenager would answer a questionnaire unless he is genuinely interested in it.

Following the example of Dominitz and Manski (1996) and Schweri, Hartog and Wolter (2008) we asked questions about expected median earnings under different schooling scenarios. Each respondent is asked to declare his expected median earnings right after graduation and 10 years later in two different cases. First, under the hypothesis that he will continue after high school to his preferred education and finish it, second in case he will stop his schooling after graduation from the school he is enrolled in at the moment of the interview.

In order to elicit their subjective earnings distribution the respondents were asked to report quantiles of this distribution. We simply asked them to report the probability that their expected wage would exceed 125% of the expected median and the probability it would fall short of 75% of their declared expected median wage – right after graduation and 10 years after it –. Since we cannot expect high school students to be familiar with the concept of the median, we gave a brief explanation of it immediately before we proceeded with these specific questions. The thresholds were given in euros rounded to the nearest 50 E, as calculated from stated medians.

Schooling expectations were also part of the enquiry. We asked if and where they wanted to continue their studies after their current school and what was the probability they would do so. To conclude, some background questions about the respondent and his/her family are asked. These questions include age, sex, father's and mother's education, school attended at the moment, school year attended and ethnicity of the respondent.

3.2 Description of Respondents

As noted, our survey was targeted at high school students of the three types of secondary school available in The Netherlands: VMBO, HAVO and VWO. The original sample was composed of 4707 individuals of which 1908 from VMBO,

1546 from HAVO and 1253 from VWO. However, for 2448 students the income variable is missing: 1409 for VMBO (74%), 586 for HAVO (38%) and 453 for VWO (36.1%). In fact, many respondents answered just the first few questions (on gender, age and education) and then quit. Also, there were 526 observations with answers violating the definition of probability: 147 for VMBO (7.7%), 197 for HAVO (12.7%) and 182 for VWO (14.5%). Also, we imposed the restriction that students should be in the last two years of their school type, and this rules out some age categories (at age 15 you cannot normally have reached grade 5 of VWO). The data loss due to this choice is minor, however, 36 observations have been eliminated for this reason: 0 for VMBO, 13 for HAVO (0.8%) and 23 for VWO (1.8%). Below, we will analyze non-response and inconsistencies, as they are important to assess the reliability of our analysis. Here we will just look at the descriptives of our resulting sample.

Our final sample is composed of 1697 individuals: 352 from VMBO, 750 from HAVO and 595 from VWO. This is smaller than we hoped for, but still compares quite favorably to Dominitz and Manski and to Schweri et al. which used a sample of 110 (71 high school and 39 college) and 252 (all college) students respectively.

Descriptive statistics for the sample we will use further on in our analyses are presented in table 1. The second column of table 1 describes the respondents from VMBO high schools. The majority of respondents are either 15 or 16, 16 being the age of graduation for a regular course of study and 15 the penultimate year of study. All the respondents declare to attend their last year of study, which for this type of school is the fourth. The figures show these individuals to come mainly from poorly educated families. Only 35.23 percent of fathers and 37.21 percent of mothers have undertaken tertiary education, many have MBO. Quite high is also the percentage of respondents who do not know the educational level accomplished by their own parents.

The fourth column illustrates the characteristics of HAVO students. If we look at column four first we see how in this case the majority of respondents are 16 and 17 years old, almost evenly divided between those two classes. Those are the ages we would expect from next-to-last and last year students from this kind of school. The family background is higher than for VMBO students. 60.67 percent of fathers and 58.67 percent of mothers have a tertiary education degree and, especially for fathers, mainly HBO or university. The sixth column refers to VWO students. The respondents are logically older than the two other schools' students. The greater part of them is either 17 or 18. There is a slight preponderance of fifth year students over the sixth year ones. Family background is constituted mainly by highly educated parents: 73.28 percent and 67.23 percent of fathers and mothers respectively have completed tertiary education.

The ethnic origin of respondents is mainly Dutch. The VMBO sample of students is slightly more ethnically mixed. It is remarkable how our sample is unbalanced in gender composition: two thirds to three quarters are girls.

4. How unreliable is an open internet survey?

4.1 Specific sensitivities

Our data gathering design might raise three sorts of concerns. In the first place, the free and voluntary participation to the survey could determine a strong self selection of our sample. Secondly, internet based surveys might be particularly affected by high attrition. This was not an issue, obviously, in laboratory data gathering. Last, as in every other survey study, the carefulness of self-reported answers can be questioned.

At first sight our procedure is easily prone to self-selection. It is realistic to believe that first coming into contact and then compiling the whole

questionnaire entails stronger interest and better information on the questions at stake relative to the average high school student. In this scenario we could expect our sample to be composed either by students who are better informed and therefore less likely to systematically declare unrealistic wage expectations, or by students more willing to assert their career choices, more self assured about it and thus more likely to unconsciously self enhance their position in the future wage distribution. If that is what happened the external validity of our findings would be severely jeopardized. Consequently, our estimates of the level of information possessed at the time of choice would be higher than for the general high school students population or the level of declared expected wages could be upwardly biased.

Internet based surveys are also intrinsically affected by missing data. Boredom, distractions and tedious compilation of numerous questions are all factors that might cause the respondent to abandon the questionnaire halfway. Since nobody can control respondents' behaviors or force them to compile the entire set of questions, the number of people abandoning after few questions can be big. If that happens and it does not happen randomly in the sample, we might believe that those that finish the questionnaire are either the most interested in the questions at stake, or those that possess a higher level of concentration. In this case the sample used for analysis would be positively selected. Expected median wages would be upwardly biased. The variance of wage distributions might also be affected. In fact, we would draw conclusions from a sample of allegedly more able students digging less deep into the ability distribution and encountering lower variance of expected wages. More contradictory would be the result of positive selection on possible bias in expected wages relative to actual wages. On one hand we might expect that more interested people are also better informed and, thus, more realistic in their expectations. On the other hand, if we selected better students they would tend to realistically locate themselves in the higher tail of the wage distribution.

The last source of concern is something that our research shares with the rest of literature exploiting survey data. Even though Dohmen, Falk, Huffman and Sunde (2005), Dominitz and Manski (1994; 1996) and Manski (2004) all

produced evidence in support of consistency of data based on surveys, it is always quite hard to assess the carefulness and attention applied by respondents during the compilation of a questionnaire. In our case, a good indicator of careless answering could be an excess of bunching towards a few suspect values.

We will address these issues to the extent possible with the information we have available. We will in turn consider non-response, bunching and inconsistencies.

4.2 Non-response

Table 1 shows the composition of the final sample we use and the sample of individuals that started to answer but gave up. We are interested in non-response on the earnings expectations, but as noted above, this non-response essentially coincides with general non-response beyond the initiating questions on gender, age and school type attended. Other questions in the survey related to risk attitude and vignettes meant to elicit preferences determining schooling choices. Most of our responders quit the survey long before our question of interest⁵. We see a remarkable discontinuity of missing values between questions thirteen and fourteen. Up to question twelve, demographical information was gathered. Question thirteen was a simple yes or no question asking whether or not the respondent receives money from his own parents. The following question (number fourteen) asked for the amount of money received and the possible use of that money. Question fourteen was answered by 1539 students less than the previous one. It is reasonable to think that many respondents could not be bothered to realistically answer to such a specific question and abandoned the survey.

Essentially, the samples of “starters” and our final sample do not differ substantially in composition by observed characteristics. The only systematic factor is gender: girls are more prone to finish the questionnaire than boys.

⁵ Precisely fifty questions before.

This does not suggest that the response on the earnings questions is very selective.

To check whether some particular characteristics correlate with the probability of leaving the sample incomplete, we run a probit regression on the probability of being or not being part of the early quitters group. The results of this regression are shown in table 2

Column I shows regression of the probability of quitting (i.e. non-response on the earnings questions) on the available controls. Some specific characteristics show a significant impact on the decision of quitting the questionnaire. Females are 11% less likely to abandon than boys, while older students and students for higher classes are more likely to do so. Particularly strong is also the impact that having a low educated father exerts; 45% more prone to quit early. Mother's education however, has no significant effect. Immigrants have no different probability of non-response, schooling level has no effect on non-response. In the second regression, we differentiate father's education by respondents' school type attended. Now we see that for VMBO and HAVO students, family background does not differentiate: the dummy coefficients are more or less equal. Only for VWO students, family background has an effect, but it is not monotonic: with father's education at medium level, non-response is markedly lower than both for higher and lower education fathers. Allowing for interaction now also brings out a marked school type effect: VMBO students have higher non-response than other students.

4.3 Bunching

In Table 3 we give one full specification of expected median earnings (starting salaries, in case of tertiary school completion). The declared expected starting wages are manifold and widely dispersed. The modal stated value, 2000 Euro per month, only attracts 14.59 percent of answers. A clear tendency of rounding at multiples of five and the presence of marked outliers at both ends of the distribution are also visible.

The bunching of expectations at multiples of integers, starting at 10 euro but then climbing to much higher base values, indicates that expectations have a rather crude nature. This is what one would anticipate: no one would think that students have expectations in multiples of one euro. The long tails of the distribution are perhaps more surprising. Some expectations are clearly far removed from actual earnings paid in the labor market. That does not rule out that there is a system in the pattern of expectations across alternatives, as we shall check below. The unrealistic values may indicate that earnings are not of prime concern in selecting an education.

4.4 Inconsistencies

Dominitz and Manski (1996) and Schweri et al. (2008) included feedback on inconsistent answers and had the survey filled in by students in a classroom situation, with a supervisor/instructor present. The laboratory setting of these studies allowed stopping the compiler once an inadmissible answer on a probability question is given and reminded the student of admissible values. Dominitz and Manski's (1996) design added further help for the respondent. Along with the help screen in case of non-admitted answers, a training screen was provided before the proper survey was administered; a help screen was always available to the respondent and review-and-revise screens allowed the respondent to apply a last check on their answers before the final value was confirmed.

We did not include any such feedback. Inconsistencies provide information on what respondents know and perceive and we did not want to compel students to conformity with rules they were not aware of. Instead, we analyse the occurrence of such inconsistencies. The only help we provided was an explanation of the concept of median right above the question regarding the expected median entry wage.

Our question on the expected median entry wage was followed in the questionnaire by two probabilistic questions. The first asked the probability

that the received entry wage would be lower than 0.75 times the median wage, while the second asked the probability of receiving 1.25 times the median wage. We observed three types of mistakes related to these two queries:

1. The respondent declared a probability exceeding 1.
2. The respondent declared two tail probabilities which sum up to more than 1;
3. The respondent assigned a value exceeding – or equal to – .50 to either the left or the right tail of the distribution.

We do not include inconsistent answers in our analysis of students' earnings expectations. This means we had to drop 526 observations, i.e. 23% of the sample. This compares favorably with Schweri et al., who interview students at tertiary level (higher vocational education). In that questionnaire, a student has to specify 20 probabilities. 65 percent of the respondents never made any error, 20 percent made one error, 11 percent made two errors and the remaining 4 percent made more than two errors (the worst was a single person with 7 errors). In Dominitz and Manski (1996) also, the share of respondents declaring a probability inconsistent with the previously reported median is comparable and in some case even higher than our case. Depending on the scenario, the values vary from a minimum of 11% to a maximum of 40%.

Table 4 summarizes the frequencies of these three types of mistakes in each subsample⁶. As we can see from this table type 1 errors are extremely rare and are almost evenly distributed in the three samples. Type 2 errors are rare as well, but some differentiation between the three samples is observable: VMBO students are more prone to this kind of error. Anyway, the bulk of errors are constituted by type 3. Once again, VMBO students show more often problems in grasping the concept of the median – as having 50% of the probability on each side – or in understanding our instructions.

⁶ The percentages of mistakes differ in table 4 from those reported on page 7. The reference population, in fact, changes in the two cases. The figures on page 7 refer to the general population of people that started the questionnaire, whilst table 4 refers only to those people that answered to the wage distribution questions.

We run a probit analysis to scrutinize whether making one of the three mistakes listed above is correlated with any observable characteristic. The dependent variable is in this case whether a respondent fell in any kind of mistake or not (type 1 and type 2 errors are too few to distinguish the three types). The dummy dependent variable will assume value 1 if the respondent answered inconsistently and 0 otherwise.

The results in Table 5 tell a clear story. Girls make fewer errors, immigrant children make more errors and students with low educated mothers make more errors. The latter effect appears to be located exclusively among HAVO students. The insignificance of the other variables is at least as interesting. In particular, we find no significant differences between types of school attended, between fathers with different levels of education or by type of intended education after graduation from the present school (Column IV). In column IV we also include the stated probability to graduate from the intended further education, and this has no significant effect either. Including future intentions interacted with school of origin as covariates can give us a rough idea of the ability of each respondent, or at least the perceived own ability. Better students, or at least students with a higher taste for education, should declare their intention to proceed with their studies either to university or HBO, while weaker students are more likely to start working after school or complete their education with an MBO degree. Future intentions might also have a heterogeneous effect depending on the school of origin. We might believe that a student declaring his will to finish education after VMBO is a different type of student and has potentially different ability to understand the concept of median than his VWO peer declaring the same intention. We would infer an effect of ability to understand our questionnaire if the coefficients would follow a monotonically increasing path from VMBO school quitters to VWO future university students. This pattern is not visible in our data. People wanting to quit their studies, continue to MBO or continue to HBO do not have a lower propensity to slip-up no matter which school they are attending at the moment

of the interview. Covariates are mainly unaffected by the inclusion of heterogeneous schooling effects. The variables that were significant before remain significant and vice versa.

Column IV also includes adds an ulterior control for ability: the self-stated probability of tertiary studies completion⁷. We see that being more or less certain on graduation has no effect on the probability of committing a mistake.

The story told by our data can comfort us for the selectivity of the sample. Our dropped observations, in fact, don't seem, from this analysis, to belong to any easily detectable disadvantaged group of students. This result gives us some confidence on the reliability of the further analysis that will exploit only those respondents able to correctly interpret the three wage distribution questions.

4.5 Assessment

Response to our survey shows a very specific pattern. Participants begin by answering the first basic questions (age, gender, school type) and then either quit or more or less fill out the entire questionnaire. Non-response is lower among girls, higher among VMBO students. Parental education has negligible effect. Students in their last school year are more inclined to answer than those in their penultimate year, which reflects growing interest in future careers. Bunching of expected median earnings is strong, but whether this is stronger than one should anticipate for this kind of information is hard to judge. Violations of the rules of probability are not more frequent in our open internet survey than in the data collected by Dominitz and Manski (1996) and by Schweri et al. (2008) for narrowly defined groups of students. Girls make significantly fewer errors than boys. Violations do not vary by type of school attended and do not vary either with other observable indicators of ability: intended continuation of education and self-assessed probability to succeed in

⁷The reason why the number of observations in this regression is inferior to the other three is, obviously, that students not willing to continue to tertiary education are here dropped since they do not declare any probability of tertiary school completion.

prolonged education. Immigrant children and children with lower educated mothers commit more errors.

In terms of relationships to observable characteristics there are no compelling reasons for serious concern on the quality and selectivity of the dataset. In terms of relationship to unobservables this is by nature harder to assess. The absence of a dominant ability effect, for given level of high school, on violations of the rules of probability may probably be extrapolated to some confidence in this respect.

5. Expectations: median earnings

5.1 Characterizing the observations

We report the expected earnings distribution of our sample of students in table 6, both for starting wages (panel A) and after 10 years (Panel B). Recall that we elicited six different earnings expectations from each respondent. Figures 1 to 4 present graphs of the data. The wage data refer to intended continued education, no matter present school attended. This implies that the graph for “To Work” is heterogeneous in completed educations.

Taking a look both at figure 1 and 2 and at table 6, we note five regularities in these data. First, for all destination groups we find high within-group variation; between-group variations are smaller. Second, expected median earnings by girls are always smaller than expected medians by boys; the gender difference increases with level of education. Third, for expectations when not continuing education, medians increase with schooling level, with a big jump towards VWO: expected medians for VMBO and HAVO are not too far apart, expected medians for VWO are about double the level for HAVO. Fourth, expected medians for continued education are generally higher than for going to work after graduation, and increase with level of continued education, but there are

exceptions. Fifth, expected earnings distributions shift upward with experience; respondents are conscious of the increasing wage profile in the life cycle of a person. If we compare panel A with panel B we see that within group variation of expected medians does not diminish with time, on the contrary. Gender differences become more marked and university ceases to be perceived as the strikingly more rewarding type of education.

5.2 Do they mirror actual data?

It would be interesting to compare individuals' expectations with actual data. However, as we have no observations on individuals' realizations, we can at best compare expectations with group means of actual data. The exercise would be of limited value. Therefore, we decided to take into consideration two other dimensions: the gender gap and the wage growth by experience. The data on actual wages – by highest degree obtained, gender and age – are taken from the Dutch Central Bureau of Statistics (CBS henceforth) survey for the year 2002.

Since absolute differences would not be informative, we interpret relative differences, namely between gender wage gap and life cycle wage profiles, as a rough, but significant, indicator of our sample's level of awareness of actual wage structure

We can not definitely rule out the possibility that disproportionate gender wage gap or unrealistic wage growth are reflecting the effect of unobservable – for us – characteristics of the respondents. Anyhow, given that we have neither positively nor negatively selected either men or women, or equally, that the type and severity of selection is the same across these two groups – an assumption that in light of our previous analysis does not seem excessively bold –, a comparison of the expected with the effective gender wage gap can provide a measure of how anchored to reality expectations are. This is what we have done and what we report in table 7.

The first column of table 7 reports the wage gap, in percentage, that women experience when compared to men with the same level of education as reported in the CBS survey for the year 2002. In the second column we show the

difference between the median female and the median male expected median wages in our survey by educational category as reported in table 5. Column III is the difference between the previous two columns. Column IV reports the coefficient of the dummy female – always negative – taken from wage regressions by educational outcome⁸. This coefficient is, thus, obtained after controlling for age, family background and ethnicity. To conclude, column 5 is the difference between column IV and I.

What is evident from column V is the systematic underestimation that, with the exception of VMBO, females make of their own future wages compared to men. Our female respondents add, on top of the already substantial gap recorded in labor market data, a handicap varying – only considering the significant coefficients – from 18.2% to 38.9%. This skepticism appears to decrease further on in the life cycle: in panel B, the difference in perceptions is closer to actual gap.

Table 8 considers the wage growth along the career path. The same structure of table 7 applies here. Column I reports wage growth as recorded in CBS (2002) data, column II the wage growth expected by the median respondent in our sample by gender, current education and future education. Column III is the difference between column II and I. Column IV reports the coefficient of the dummy 10 years after in a wage regression pooled by scenario and stratified by gender and declared future education⁹. To conclude, column V is the difference between column IV and III.

At a first glance, what should immediately strike us is the low adherence of expected wage dynamics to recorded wage dynamics. The sign of mistakes

⁸ The estimated regression takes the form: $\ln W_c = \alpha_c + \beta_c F + \gamma_c X_{ic} + \varepsilon_c$ where W indicates the logarithm of expected wages, F a dummy taking value 1 for females, X a vector of controls for family background, age and ethnicity ε a disturbance term and c the specific educational category that the respondent declared to have the intention to be part of. The coefficient shown in table 7 is β_c

⁹ The exact wage equation estimated here is:

$$\ln W_{c,g} = \alpha_{c,g} + \beta_{c,g} 10AFT + \chi_{c,g} STOP + \delta_{c,g} 10AFT_STOP + \gamma_{c,g} X_{i,c,g} + \varepsilon_{c,g}.$$

Where W indicates the expected entry wage, $10AFT$, $STOP$ and $10AFT_STOP$ the scenario dummies X a vector of controls for family background, age and ethnicity, ε a disturbance term, c the specific educational category that the respondent declared to have the intention to be part of and g the gender. The coefficient reported in column IV is $\beta_{c,g}$.

changes without an apparent pattern; it does not follow a monotonic path from lower to higher educational category or vice versa and is equally substantial for both girls and boys. It is also quite remarkable that controlling for family background, age and ethnicity equalizes the expected growth: the coefficients of column IV are all pretty close to 0.6.

Always keeping in mind that our type of questions were not designed for a specific investigation of awareness possessed, and that we cannot rule out the possibility that private information plays a role in the formation of expectations, the two abovementioned comparisons could shed some doubt on the overall quality of knowledge on real labor market dynamics and conditions in our particular high school students sample. The gap in wage expectation between men and women is substantially larger than the gap in actual mean wages between men and women. Growth in expected wages over the future 10 years differs substantially from actual growth in mean wages, but there is no systematic under or overestimation detected.

5.3 Explaining expected medians

In Tables 9, 10 and 11, we report results of estimated wage regressions for expected medians, for VMBO, HAVO and VWO students respectively. Recall that the respondents were asked to declare median wages (and wage dispersions) in four different scenarios:

1. Entry wage in case of secondary education degree only;
2. Entry wage given tertiary education degree;
3. Wage after ten years of work in case of secondary education degree only;
4. Wage after ten years of work in case of tertiary education degree.

We estimated four different equations for each different scenario.

Tables 9, 10 and 11 show some clear results. First, girls expect between 24 and 46% lower medians than boys. The effect is present in all scenario's for all school types. Second, there is no clear systematic effect of family background on expected medians. Mother's education has no effect at all (among 24
20

coefficients, only one is significant), Father's education has no effect at all for HAVO students, low education reduces expected starting salaries for VMBO students, low and median education increases expected salaries in some scenario's for VWO. Whereas the empirical literature generally finds a positive effect of better family backgrounds on earnings, we do not observe this pattern in our regressions. Third, ethnicity ("immigrant") has no effect on expected medians. Fourth, the evidence on self-selection in expected benefits is mixed at best. Among VMBO students, there is nowhere a significant difference in expected wages between those who go to work and those who continue education. Those who do continue expect lower salaries in all scenarios than those who do not continue. This is probably not what one would have predicted: those who do continue are not the better students in terms of their expected wages under all scenarios. However, choice is supposed to respond to the mark-up on tertiary education and this is what we will consider in section 7.

Among HAVO students, again no wage difference between those who continue school and those who go to work is statistically significant. Only among VWO students we do find significant expected wage differences between those who go to university and those who don't. Those who do not intend to attend university expect close to 20% lower wages from university degree than those who do, whereas they also expect a lower wage if they go straight to work after completing VWO – although not significantly so. Now, the expected probability of completion is also significant: students with higher expected probability to complete university have lower expected opportunity costs from attending university. In all other cases, self-rated probability to complete a tertiary education has no significant effect on any expected wage.

6. Expectations: Quantifying risk

6.1 Measuring risk and skewness

In the literature on the Risk Augmented Mincer earnings function (King, 1974; Hartog and Vijverberg 2007; Hartog, Plug, Serrano and Vieira, 2003; Diaz-

Serrano, Hartog and Nielsen, 2003) it is common to characterize risk as the dispersion of wages around the mean for a particular educational group to which the individual belongs. However, this approach has been criticized for drawing conclusions from ex-post wage realizations to measure ex-ante wage risk (Cunha, Heckman and Navarro, 2005). The variance found in wage data is not necessarily imputable to uncertainty. Heterogeneity may play an important role as well. Individuals may be better informed than researchers about their own ability and they will use this superior knowledge when forming their wage expectations. Neglecting heterogeneity will cause an overestimation of risk (Cunha et al. 2005). Our strategy to overcome this criticism is that of asking students directly about their expectations as in Dominitz and Manski (1996) and Schweri et al. (2008).

As explained in Section 3.1, for each respondent we have three points of his expected earnings distribution: 1.00, 0.75 and 1.25 times the median of their wage, at entry and ten years later. If we want to calculate variance and skewness from this information we need to impose some distributional assumptions. Dominitz and Manski fitted a log-normal distribution. Since we doubt that such a distribution would reflect correctly the distribution that respondents had in mind, we will not impose any distributional assumption. On the contrary, for our measure of risk we will follow the steps of Schweri et al. (2008) and specify an alternative indicator for variance and skewness.

The three points of the distribution that we know make it possible to divide the probability density function in four intervals: $(0-0.75*m]$ $(0.75*m-m]$ $(m-1.25*m]$ $(1.25*m-\infty)$, m indicating the declared expected median value. We denote the probability masses lying in the four intervals as A, B, C and D respectively. By the definition of median $A+B=C+D=0.5$. A measure of variance could thus be given by the share of probability assigned to the outer intervals of the distribution: $v=(A+D)$. v lies between 0 and 1 and it will be our variance indicator.

The skewness coefficient exploits a similar stratagem. In fact, looking at the asymmetry of the distribution, we can obtain our skewness coefficient s as

given by: $s=2(D-A)$ and it varies between -1 and 1. A positive value indicates positive skewness, the opposite holds for a negative value.

6.2 Characterizing variance and skewness

In figures 5 to 12 we provide a visual description of the distributions of the variance and skewness coefficients across individuals in our sample. The distributions are defined on a fixed interval for a total of 1693 cases¹⁰: 1207 girls and 486 boys.

What emerges from a visual analysis is the high risk perceived by many. In fact, even though the distribution is almost symmetrical around 0.5, many of the individuals in our sample foresee a variance coefficient equal to one, the maximum possible (admittedly, this implies a somewhat weird distribution, with no probability mass between 0.75m and 1.25m).¹¹

It is also remarkable to note that type of education doesn't affect the level of risk: the four distributions are very similar to another. Girls perceive higher wage risk and both boys and girls expect wage variance to decrease with time. For the skewness coefficient the story is somehow simpler. For entry wages, the distributions are almost symmetrical around zero, with most of observations concentrated around the mean. The high frequency of a symmetric distribution (skew equal to zero) is also remarkable.

No significant gender or future school type differentiation is traceable. What is striking is the noteworthy translation to the right that this distribution shows after 10 years from entry into the labor market. Once again the same path emerges for boys and girls and for whichever type of degree earned.

6.3 Explaining variance and skew

In tables 12-15, we present regression results for variance and skewness, stratified by type of education as before. Barely any of our explanatory

¹⁰ The cases for females divide as follows: 198 for MBO, 611 for HBO, 378 for University and 20 go to work. For males the specific figures are: 85 for MBO, 209 for HBO, 175 for University and 17 go to work.

¹¹ 328 respondents reported a peculiar wage distribution for which the variance coefficient equals 0 and the skewness coefficient equals 1.

coefficients is statistically significant. There is some inclination for women to expect higher variance in their future earnings distributions notably for VMBO and VWO students. For VWO students the magnitude does not depend on scenario, while for VMBO students it is higher for tertiary education. Joint with higher risk aversion among women, this would count as a deterrent on continued education. The overriding impression from the tables is lack of systematic influences: neither ethnicity, nor family background, nor “ability” (intended further studies, probability to complete) have any explanatory power.

There are three potential interpretations. The results may reflect purely private information (expectations only depend on unobservables), complete randomness (there is no useful information in the answers), or they may reflect variation in perceptions of actual labor outcomes. The first interpretation is not very convincing. If individuals were to use unobservable private information, one would expect them also to use information on their abilities, such as intended further studies and probability to graduate. The third interpretation can not be tested with the present dataset, but finds some support in analysis of comparable Swiss data (Schwery et al., 2008). The second interpretation may be tested by linking perceived risk to actual behavior; one might, for example, relate educational choice to perceived risk in combination with information on individuals’ risk attitude. This would also require additional information.

7. Expectations: returns to education

7.1 Eliciting perceived returns to education

As noted before our survey elicited wage expectations in four different scenarios¹². This particular feature of the questionnaire allow us to calculate the perceived returns to education for subjects declaring expected earnings,

¹² See section 5.3.

evolution in time given the accomplishment of a particular education and the counterfactual in case of non accomplishment.

The internal rate of return to education is the rate of discount ρ equating the present values of lifetime earnings for the two different scenarios – accomplishment or non accomplishment of desired tertiary education. We can obtain this interest rate by solving the equation:

$$\int_{t_0}^{t_1} y_{1,t_0} e^{-\rho t} dt + \int_{t_1}^T y_{2,t_1} e^{-\rho t} dt = \int_{t_0+s_i}^{t_1+s_i} y_{3,t_0+s_i} e^{-\rho t} dt + \int_{t_1+s_i}^T y_{4,t_1+s_i} e^{-\rho t} dt \quad (1.1)$$

The left and right hand side of the equation describe the life cycle earnings of an individual in case of no further education and in case of accomplished tertiary education respectively. y_1 indicates the expected earnings at time of entry in the labor market (t_0), in our case this time is equal to zero. The second period of life starts at t_1 , which as we have said, for us corresponds to ten years after leaving school, and y_2 are the expected earnings during this period until retirement age which occurs at T. We assume retirement age to occur forty years after entry in the labor market without extended education. On the right hand side of equation 1.1 s is the additional years of schooling necessary to obtain the diploma each respondent declared to be willing to achieve. Expected earnings in case of tertiary education completion are described by y_3 for entry wages and y_4 for wages starting ten years after graduation.

We possess information about the four different y directly provided by the respondents. It is then possible to plug in these values in equation 1.1 and numerically solve it, obtaining our parameter of interest ρ ¹³.

Some descriptive statistics of the distribution of the perceived rate of return to education in our sample are provided in table 18. As we can see it does not differ at all between genders. Both females and males show a median perceived rate of return equal to 0.17 with the same inter-quantile range (IQR) as well (0.18). HBO students present the lower expected return from their educational

¹³ For this purpose we used the software matlab inputting the possessed data and exploiting the automatic equation solver provided in this software.

investment (0.13) for both men and women, whilst future university students the highest with 0.23. Future HBO students also show the lowest IQR – 0.15 for women and 0.17 for men – while future workers the highest – 0.44 for female and 0.25 for men.

7.2 Characterizing perceived returns to education

Figures 13 and 14 report the distribution of perceived returns to education stratified by future education for males and females respectively. We are not differentiating here by current education, even though we are aware that such a distinction might be useful, because further fragmentation of our sample would make this visual analysis somehow less significant. In the next section, performing regression analysis on the determinants of expected returns to education, we will control for both present and future education.

From these distributions it seems clear that no dramatic differences exist between the three types of schooling for both men and women. Two characterizing factors emerge. First, expected returns for future HBO students are slightly lower with a fatter left tail of the distribution. Second, future HBO's students present a lower dispersion around the median. Both these facts are common to boys and girls.

7.3 Explaining perceived returns to education

Table 19 presents simple OLS regressions on our usual demographics, family background characteristics and proxies for ability – intention to continue and self-stated probability of completion – on the perceived returns to education stratified by respondents' current school as obtained from equation 1.1.

Just like for the variance and skewness regression, also in this case the observables at hand have a really low explanatory power. Ethnicity and family background have no systematic effect on expected returns. Only the probability of completion shows a positive contribution to the perceived benefit of education for both VMBO and VWO students. "Better" students expect a higher rate of return. In the case of VWO students also the dummy "not university" has a strong negative effect on the dependent variable. This is an

interesting corroboration of self-selection: students intending to continue to university have substantially higher rate of return than those who do not intend to continue.

One thing that clearly emerges from this simple regression is the lack of any gender differential in expected returns. Even though women expect to earn less in their future careers, they do not perceive their investment to be less profitable than men.

8. Testing the risk augmented Mincer earnings function.

The Risk Augmented Mincer earnings regression takes the form:

$$\ln w_i = X_i\beta_1 + R_i\beta_2 + K_i\beta_3 + \varepsilon_i. \quad (1.2)$$

In this equation R_i and K_i measure risk and skewness, respectively, while X_i represents the usual vector of personal characteristics, including schooling. If our sample is composed by risk averse people, we would expect R to positively affect wages. That is because risk averse individuals will request, through higher wages, compensation for the uncertainty they dislike. The effect of K should, instead, be negative. Workers appreciate a wage distribution skewed to the right. They would be ready to pay for this preference by accepting lower wages. These effects are commonly observed using both realization data (Hartog and Vijverberg, 2007; Hartog et al. 2003; Diaz-Serrano et al. 2003) and expectations data (Schweri et al. 2008).

Estimated regressions with the present dataset do not support the predictions. The variance coefficient has statistically insignificant effect in every subsample and in every scenario. The skewness coefficient shows a significant, but positive effect on expected median wage. We have no convincing explanation for this stark contrast between the results on the Swiss university students and our dataset of high school students. Perhaps there are differences between high school students and university students that do provide an explanation:

university students, being older, may have a better understanding of the labor market. Note however, that ability to understand the workings of the labor market or to understand the survey questions is not a strong candidate for explanation: students set to go to university are included in our sample.

8. Conclusion

We started this work with some questions in mind: the level of knowledge of Dutch high school students on labor market conditions both in terms of wages paid and in terms of variability of wages; the dynamics of information accumulation; the presence of perceived discrimination that by itself might influence realized wages via lower benefits from effort. Under a methodological point of view, the present paper being the first in this field dealing with an internet based survey, we had to examine whether such a methodology is suitable for our intentions.

Our internet based survey allowed us to easily and quickly come into contact with a vast audience of respondents blowing up the number of observations available. We did not find compelling evidence that our survey is evidently biased, e.g.: towards abler or more interested students. The number of erroneous answers – as violations of probability theory – was comparable with earlier, laboratory led, experiments, but by omitting feedback on the errors we did loose a large number of observations. Mostly, our survey was afflicted by a high number of early quitters. The population to which our survey was addressed might be particularly prone to leaving surveys unaccomplished. Coming into contact with a questionnaire via internet and probably at home with no supervision deprived this questionnaire of some importance in the eyes of teenagers. It is not unlikely that many of them started it only to kill some time and easily stopped once bored or once they realized that some concentration was required. This problem would, presumably, afflict every internet based survey addressed to such a young population and is not specific

to the method of obtaining information on wage expectations. Probably a more accurate design, length and phrasing could contribute in minimizing the number of early quitters.

The questionnaire generated many interesting conclusions. As in earlier work, the direct survey method points to wide dispersion of expected median earnings among individuals, but few systematic influences can be established. The same holds for variance and skewness. Students allow for wide dispersion in the wages distributions associated with a particular education scenario and dispersions vary strongly between individuals. But again, there are very few systematic patterns that can explain the differences between individuals. In particular, the absence of influence of family background and ability indicators is at variance with what self-selection models based on private information would lead one to anticipate. Surprisingly enough they do not seem to demand any type of compensation for this risk.

A most remarkable result in our analysis is the systematic lower wage expectations that females have for themselves. However, this is a systematic effect in all expectations and by consequence, anticipated rates of return to education do not vary by gender. The offspring of originally non-Dutch families do not share lower wages than natives.

Rates of returns implicit in students' expectations show wide variation across individuals with a mean of 0.17 and an interquartile range of 0.18 for both genders. Anticipated rates of return differ by education but not monotonically with education level.

Some evidence is in line with self-selection models. Perceived probabilities to complete an extended education increase the perceived rate of return, for VMBO and VWO students, not for HAVO students. And secondary school students that do not intend to continue to university foresee a lower return to university training than students that do intend to continue. But in expected earnings levels we do not see patterns in line with self-selection; among VMBO students and among HAVO students, wage expectations do not differ by intention to continue or not.

In all, we think the method of Dominitz and Manski can certainly be applied in open internet surveys. But much more work is needed to unravel the precise nature of wage expectations.

References

- J. Betts (1996)**, “What Do Students Know about Wages? Evidence from a Survey of Undergraduates.” *The Journal of Human Resources*, 31(1), pp. 27-56.
- F. Blau and M. Ferber (1991)**, “Career Plans and Expectations of Young Women and Men.” *The Journal of Human Resources*, 26, pp. 581-607.
- G.J. Borjas (2008)**, *Labor Economics*. New York, N.Y.: McGraw-Hill.
- F. Cunha, J. Heckman and S. Navarro (2005)**, “Separating Uncertainty from Heterogeneity in Life Cycle Earnings.” *Oxford Economic Papers*, 57(2), pp. 191-261.
- L. Diaz-Serrano, J. Hartog and H.S. Nielsen (2008)**, “Compensating Wage Differentials for Schooling Risk in Denmark.” *Scandinavian Journal of Economics* (forthcoming).
- T. Dohmen, A. Falk, D. Huffman and U. Sunde (2005)**, Individual Risk Attitudes: New Evidence from a Large Representative Experimentally Validated Survey, IZA Discussion Paper, 1730.
- J. Dominitz and C.F. Manski (1996)**, “Eliciting Student Expectations of the Returns to Schooling.” *The Journal of Human Resources*, 31(1), pp. 1-26.
- J. Dominitz and C.F. Manski (1994)**, “Using Expectation Data to Study Subjective Income Expectations.” Social System Research Institute Working Paper 9444, University of Wisconsin-Madison.
- J. Hartog, E. Plug, L.D. Serrano and J. Vieira (2003)**, “Risk Compensation in Wages – a Replication.” *Empirical Economics*, 28 pp. 639-647.
- J. Hartog and W. Vijverberg (2007)**, “On Compensation for Risk Aversion and Skewness Affection in Wages.” *Labour Economics*, 14(6) pp. 938-956.
- R.A. Hedley (1986)**, “Everybody but Me: Self-Other Referents in Social Research.” *Sociological Quarterly*, 56, pp. 245-258.
- A. G. King (1974)**, “Occupational Choice, Risk Aversion and Wealth.” *Industrial and Labor Relations Review*, pp. 586-596.

- C. F. Manski (2004)**, "Measuring Expectations." *Econometrica*, 72(5), pp. 1329-1376.
- J. Schweri, J. Hartog, S.C. Wolter (2008)**, Do Students Expect Compensation for Risk? Working paper Amsterdam/Bern.
- H.L. Smith and B. Powell (1990)**, "Great Expectations: Variation in Income Expectations Among College Seniors." *Sociology of Education*, 63(3), pp. 194-207.
- O. Svenson (1981)**, "Are We All Less Risky and More Skilful than Our Fellow Drivers?" *Acta Psychologica*, 47, pp. 143-148.
- S.C. Wolter (2000)**, "Wage Expectations: A Comparison of Swiss and US Students." *Kylos*, 53, pp. 51-69.
- B. Zafar (2011)**, "How Do College Students Form Expectations?" *Journal of Labor Economics*, 29, pp. 301-348.

Table 1. Characteristics of respondents (in percentages)

	VMBO		HAVO		VWO	
	Original sample	Analyzed sample	Original sample	Analyzed sample	Original sample	Analyzed sample
N	1908	352	1546	750	1253	595
Female	57.49	63.92	66.75	74.93	65.68	71.13
Age						
14	13.84	2.27	1.36	0	0.08	0
15	33.60	25.85	15.98	15.87	3.35	0
16	31.76	51.99	34.15	35.73	21.71	23.87
17	8.56	15.06	30.72	34.00	34.56	42.69
18 or more	8.96	4.83	17.59	14.40	40.06	33.45
Year in school						
4th year	48.68	100	50.45	49.47	0	0
5th year	2.22	0	49.55	50.53	49.64	52.44
6th year	2.80	0	0	0	50.36	47.56
Father's education						
Primary education or similar	8.63	7.39	4.28	4.00	3.61	2.69
VMBO or similar	22.97	26.14	16.93	15.73	10.82	12.94
HAVO/VWO or similar	8.4	7.67	9.71	8.13	7.97	6.72
MBO or similar	17.93	20.17	22.96	24.00	19.38	21.85
HBO/University or similar	16.02	15.06	34.87	36.67	52.85	51.43
Doesn't know	22.91	20.45	9.84	9.87	4.87	4.20
Not applicable (no father)	3.14	3.13	1.41	1.60	0.50	0.17
Mother's education						
Primary education or similar	8.24	6.25	4.28	2.93	3.52	2.52
VMBO or similar	25.77	26.99	19.61	18.67	13.93	13.61
HAVO/VWO or similar	10.14	7.39	11.71	10.80	11.33	11.93
MBO or similar	21.23	25.85	29.12	32.27	22.82	23.87
HBO/University or similar	10.76	11.36	26.17	26.40	43.29	43.36
Doesn't know	22.13	20.74	8.23	8.27	4.87	4.37
Not applicable (no mother)	1.74	1.42	0.87	0.67	0.25	0.34
Ethnicity						
Dutch	88.7	92.05	93.56	95.07	93.45	95.63
Moroccan	2.87	1.42	1.34	1.60	0.84	0.67
Turkish	2.27	2.02	1.33	1.32	0.92	0.17
Surinamer	3.04	1.14	2.15	2.13	2.35	1.34
Antillean	1.57	1.14	1.01	0.93	0.67	0.17
Belgian	1.57	1.14	0.74	0.53	0.92	0.67
Other	8.77	6.82	5.7	5.20	7.06	6.55

Note: The respondents could select more than one option for the ethnicity question therefore the percentages can sum up to more than 100.

Table 2. Probit regression for non responses (marginal effects)

	I	II
Female	-0.115***(0.016)	-0.126***(0.016)
Age	0.049***(0.008)	-0.015*(0.007)
School year penultimate	0.273***(0.017)	0.201***(0.016)
Immigrant	0.015(0.023)	0.034(0.022)
Father's education low	0.456***(0.019)	
Father's education medium	0.053*(0.021)	
Mother's education low	-0.010(0.038)	0.003(0.037)
Mother's education medium	-0.038(0.022)	-0.022(0.021)
VMBO	-0.013(0.038)	
HAVO	-0.016(0.021)	
Father education low: VMBO		0.208***(0.039)
Father education medium: VMBO		0.186***(0.026)
Father education high: VMBO		0.169***(0.023)
Father education low: HAVO		-0.182***(0.066)
Father education medium: HAVO		-0.142***(0.031)
Father education high: HAVO		-0.193***(0.021)
Father education low: VWO		-0.093(0.083)
Father education medium: VWO		-0.240***(0.036)
PseudoR2	0.143	0.103
N	4.707	4.707

Note: robust standard errors in parenthesis. ***/**/* indicates significance levels at 10%/5%/1%. Reference group: Male, last year of high school, Dutch, VWO, father's education high, mother's education high, father's education high: VWO.

Table 3. Frequencies of values for elicited medians – entry wages

Reported value	Freq.	Percent	Reported value	Freq.	Percent	Reported value	Freq.	Percent
0	16	0.93	1120	1	0.06	3000	81	4.73
1	4	0.23	1150	2	0.12	3100	1	0.06
3	1	0.06	1195	1	0.06	3200	1	0.06
10	2	0.12	1200	102	5.95	3250	1	0.06
17	1	0.06	1250	14	0.82	3400	1	0.06
20	1	0.06	1300	23	1.34	3500	21	1.23
50	10	0.58	1350	7	0.41	3600	1	0.06
60	2	0.12	1400	20	1.17	4000	22	1.28
70	1	0.06	1450	1	0.06	4028	1	0.06
75	1	0.06	1500	220	12.84	4500	3	0.18
80	2	0.12	1519	1	0.06	5000	30	1.75
90	1	0.06	1530	1	0.06	5500	1	0.06
100	19	1.11	1550	2	0.12	6000	3	0.18
115	1	0.06	1600	26	1.52	6500	1	0.06
120	2	0.12	1700	33	1.93	7000	3	0.18
130	1	0.06	1750	11	0.64	7500	2	0.12
136	1	0.06	1800	56	3.27	7890	1	0.06
150	9	0.53	1843	1	0.06	8000	2	0.12
200	26	1.52	1900	8	0.47	9000	1	0.06
250	7	0.41	1950	1	0.06	10000	7	0.41
251	1	0.06	1981	1	0.06	11000	1	0.06
255	1	0.06	2000	250	14.59	12000	1	0.06
300	27	1.58	2068	1	0.06	13000	2	0.12
320	1	0.06	2100	8	0.47	15000	1	0.06
350	5	0.29	2125	1	0.06	16000	1	0.06
400	9	0.53	2166	1	0.06	20000	1	0.06
440	1	0.06	2200	18	1.05	70000	1	0.06
450	4	0.23	2250	2	0.12	100000	1	0.06
500	73	4.26	2300	8	0.47			
550	2	0.12	2350	1	0.06	Total	1,714	100
560	2	0.12	2360	1	0.06			
600	20	1.17	2400	11	0.64			
650	2	0.12	2500	90	5.25			
700	20	1.17	2510	1	0.06			
750	11	0.64	2600	1	0.06			
780	1	0.06	2700	3	0.18			
800	58	3.38	2750	3	0.18			
850	1	0.06	2800	3	0.18			
900	22	1.28	2810	1	0.06			
1000	190	11.09	2848	1	0.06			
1100	22	1.28	2916	1	0.06			
1111	1	0.06	2996	1	0.06			

Table 4. Mistakes by Type and School (in percentages)

	VMBO	HAVO	VWO
Declared probability exceeds 1.00	1.4	1.1	0.9
Sum of probabilities exceeds 1.00	5.4	2.9	3.0
Probability in tail below/above median exceeds .50	27.8	19.0	21.9
Total	29.4	20.5	22.7
N	499	960	800

Table 5. Probit regression for errors (marginal effects)

	I	II	III	IV
Female	-0.068***(0.020)	-0.065**(0.020)	-0.068***(0.020)	-0.065**(0.021)
Age	-0.010(0.013)	-0.008(0.013)	-0.010(0.011)	-0.008(0.013)
School year	0.010(0.025)	0.010(0.025)	0.007(0.023)	0.012(0.025)
Immigrant	0.111***(0.030)	0.104***(0.031)	0.112***(0.030)	0.105***(0.031)
VMBO	0.039(0.036)			
HAVO	-0.030(0.022)			
Father's education low	0.024(0.045)	0.029(0.045)	0.023(0.045)	0.040(0.047)
Father's education medium	0.004(0.025)	0.005(0.025)	0.002(0.025)	0.005(0.025)
Mother's education low	0.164**(0.053)	0.165**(0.053)		0.161**(0.054)
Mother's education medium	0.028(0.024)	0.028(0.024)		0.026(0.024)
From VMBO to work/MBO		0.029(0.038)		0.027(0.039)
From VMBO to HBO/WO		0.086(0.057)		0.086(0.057)
From HAVO/VWO to work/MBO		0.022(0.061)		-0.005(0.076)
From HAVO to HBO		-0.040(0.024)		-0.040(0.024)
From HAVO to university		0.048(0.054)		0.049(0.054)
From VWO to HBO		-0.012(0.038)		-0.012(0.038)
Mother education low: VMBO			0.147(0.085)	
Mother education medium: VMBO			0.088(0.047)	
Mother education high: VMBO			0.057(0.042)	
Mother education low: HAVO			0.217**(0.083)	
Mother education medium: HAVO			-0.004(0.033)	
Mother education high: HAVO			-0.017(0.024)	
Mother education low: VWO			0.144(0.104)	
Mother education medium: VWO			0.035(0.038)	
Probability of completion				0.000(0.000)
PseudoR2	0.027	0.029	0.027	0.032
N	2.207	2.207	2.207	2.154

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%. Reference group: Male, last year of high school, Dutch, VWO, father's education high, mother's education high, from VWO to university, mother education high: VWO.

Table 6. Earnings Expectations

A. Quantiles of expected earnings after graduation												
Respondent Group	Stop Studying after current school			Proceeding to MBO			Proceeding to HBO			Proceeding to University		
	Empirical Quantile			Empirical Quantile			Empirical Quantile			Empirical Quantile		
	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
Female VMBO	<i>200</i>	<i>800</i>	<i>1200</i>	200	1000	2000	500	1500	2500	/	/	/
Male VMBO	<i>650</i>	<i>950</i>	<i>2000</i>	500	1500	2500	500	1800	3000	<i>300</i>	<i>3000</i>	<i>5500</i>
Female HAVO	<i>50</i>	<i>1000</i>	<i>1500</i>	700	1500	2000	320	1400	2500	750	1900	5000
Male HAVO	<i>800</i>	<i>1050</i>	<i>1500</i>	<i>500</i>	<i>2000</i>	<i>3000</i>	750	1800	3000	1000	2500	7000
Female VWO	<i>1500</i>	<i>2100</i>	<i>2700</i>	/	/	/	500	1500	2500	600	1650	3000
Male VWO	<i>800</i>	<i>2200</i>	<i>5000</i>	/	/	/	900	2000	2500	1000	2000	3500

B. Quantiles of expected earnings 10 years after graduation												
Respondent Group	Stop Studying after current school			Proceeding to MBO			Proceeding to HBO			Proceeding to University		
	Empirical Quantile			Empirical Quantile			Empirical Quantile			Empirical Quantile		
	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
Female VMBO	<i>1000</i>	<i>1200</i>	<i>1800</i>	450	1250	2500	788	2000	4000	/	/	/
Male VMBO	<i>650</i>	<i>1850</i>	<i>3000</i>	1000	2000	4000	600	2200	10000	<i>1000</i>	<i>2900</i>	<i>8000</i>
Female HAVO	<i>300</i>	<i>1500</i>	<i>2500</i>	1000	1550	3000	700	2000	4000	1340	2200	5000
Male HAVO	<i>1500</i>	<i>2250</i>	<i>3000</i>	<i>1000</i>	<i>3500</i>	<i>4000</i>	1350	2500	6000	1400	4000	12000
Female VWO	<i>3500</i>	<i>4250</i>	<i>5000</i>	/	/	/	850	2000	4000	1000	2500	5000
Male VWO	<i>3000</i>	<i>5000</i>	<i>10000</i>	/	/	/	1300	2500	8000	1800	3200	7000

Note: there are no observations of VMBO females respondents declaring the intention to proceed to University. It is not possible for VWO students to proceed towards MBO. In italics those cells for which the number of observations is less than 20

Table 7. Gender wage gap, recorded and perceived

A. Entry wage					
	CBS data recorded gap	Our sample perceived gap, crude data	Difference	Our sample perceived gap, regression data	Difference
	I	II	III	IV	V
VMBO	0.33	0.158	-0.172	0.464	0.134
HAVO	0.31	0.48	0.262	0.128	-0.182
VWO	0.31	0.46	0.264	0.793	0.483
MBO	0.183	0.243	0.06	0.371**	0.188
HBO	-0.166	0.146	0.313	0.223**	0.389
University	0.109	0.158	0.049	0.291**	0.182
B. Wage 10 years after graduation					
	CBS data recorded gap	Our sample perceived gap, crude data	Difference	Our sample perceived gap, regression data	Difference
	I	II	III	IV	V
VMBO	0.393	0.351	-0.042	0.609	0.216
HAVO	0.213	0.334	0.121	0.034	-0.179
VWO	0.213	0.15	-0.63	0.403	0.190
MBO	0.325	0.491	0.166	0.302**	-0.023
HBO	0.254	0.167	-0.087	0.403***	0.149
University	0.336	0.302	-0.034	0.399***	0.063

Note: */**/** indicates significance levels at 10%/5%/1%. CBS groups HAVO and VWO graduates in one category. Column I is obtained from the ratio: $W_{F,c} / W_{M,c}$ in the CBS data. Column II is obtained from the ratio: $W_{F,c} / W_{M,c}$ in our data. *F* and *M* indicate females and males respectively and *c* the educational category.

Table 8. Wage growth, recorded and perceived

	CBS data recorded growth	Our sample perceived growth, crude data	Difference (I-II)	Our sample perceived growth, regression data	Difference (I-IV)
	(I)	(II)	(III)	(IV)	(V)
Females VMBO	0.843	0.500	-0.343	0.519**	-0.324
Females HAVO	1.204	0.500	-0.704	0.455	-0.749
Females VWO	1.204	1.023	-0.181	0.670**	-0.534
Females MBO	0.224	0.120	-0.104	0.608***	0.384
Females HBO	0.194	0.363	0.169	0.573***	0.379
Females University	0.060	0.324	0.264	0.700***	0.64
Males VMBO	1.038	0.947	-0.091	0.099	-0.939
Males HAVO	0.942	1.143	0.201	0.639**	-0.303
Males VWO	0.942	1.273	0.331	0.592	-0.35
Males MBO	0.482	0.571	0.089	0.517***	0.035
Males HBO	0.920	0.286	-0.634	0.720***	-0.2
Males University	0.424	0.347	-0.077	0.838***	0.414

Note: */**/** indicates significance levels at 10%/5%/1%. CBS groups HAVO and VWO graduates in one category. The coefficient of column II are calculated as the percentage increase of wages in 10 years:

$$\frac{W_{10} - W_0}{W_0}$$

Table 9. Wage regression VMBO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.414** (0.133)	-0.279 (0.150)	-0.382** (0.123)	-0.264* (0.124)
Age	0.091 (0.084)	0.002 (0.112)	-0.045 (0.091)	-0.073 (0.074)
Immigrant	0.272 (0.216)	0.274 (0.198)	0.276 (0.152)	0.148 (0.235)
Father's education low	0.002 (0.154)	0.059 (0.171)	-0.147 (0.189)	-0.143 (0.143)
Father's education medium	-0.512** (0.190)	-0.044 (0.209)	-0.405* (0.159)	-0.269 (0.163)
Mother's education low	-0.172 (0.231)	-0.329 (0.235)	-0.002 (0.210)	-0.146 (0.194)
Mother's education medium	0.191 (0.173)	0.063 (0.207)	0.302* (0.143)	0.171 (0.153)
Continue: MBO	-0.541 (0.375)	-0.711 (0.500)	-0.377 (0.268)	-0.595 (0.320)
Continue: HBO	-0.394 (0.383)	-0.455 (0.508)	0.008 (0.292)	-0.347 (0.348)
Probability to complete	0.000 (0.004)	0.003 (0.005)	0.003 (0.003)	0.005 (0.003)
Constant	6.760*** (0.506)	7.453*** (0.688)	7.333*** (0.537)	7.927*** (0.447)
R2	0.091	0.041	0.089	0.063
N	300	286	330	314

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%. Reference group: Male, Dutch, father's education high, mother's education high, not proceeding to tertiary education

Table 10. Wage regression HAVO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.239** (0.088)	-0.306*** (0.085)	-0.246** (0.081)	-0.460*** (0.071)
Age	-0.030 (0.053)	-0.036 (0.054)	0.058 (0.051)	0.010 (0.048)
Immigrant	0.125 (0.147)	0.292* (0.139)	-0.081 (0.146)	0.144 (0.104)
Father's education low	0.172 (0.216)	-0.021 (0.113)	-0.073 (0.123)	0.022 (0.167)
Father's education medium	-0.027 (0.104)	-0.001 (0.089)	-0.026 (0.092)	-0.112 (0.089)
Mother's education low	-0.047 (0.197)	0.129 (0.115)	0.125 (0.131)	-0.175 (0.136)
Mother's education medium	0.118 (0.088)	0.111 (0.080)	0.073 (0.078)	0.156 (0.082)
Continue: HBO	-0.040 (0.275)	0.179 (0.301)	0.104 (0.276)	-0.093 (0.284)
Continue: University	-0.149 (0.340)	0.035 (0.344)	0.481 (0.328)	0.076 (0.310)
Next-to-last year student	0.014 (0.100)	-0.004 (0.106)	-0.059 (0.081)	-0.094 (0.088)
Probability to Complete	-0.001 (0.003)	-0.002 (0.003)	0.001 (0.003)	0.002 (0.003)
Constant	7.055*** (0.365)	7.518*** (0.365)	6.869*** (0.361)	7.832*** (0.336)
R2	0.020	0.035	0.036	0.083
N	693	670	742	717

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%. Reference group: Male, Dutch, father's education high, mother's education high, MBO or not proceeding to tertiary education.

Table 11. Wage regression VWO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.252** (0.083)	-0.347*** (0.089)	-0.252*** (0.076)	-0.395*** (0.080)
Age	0.059 (0.056)	0.119* (0.059)	0.047 (0.055)	0.018 (0.052)
Immigrant	0.086 (0.096)	0.064 (0.088)	0.101 (0.071)	0.086 (0.083)
Father's education low	0.266 (0.179)	0.377** (0.142)	0.441** (0.162)	0.211 (0.147)
Father's education medium	0.189** (0.072)	0.164* (0.068)	-0.017 (0.071)	-0.044 (0.072)
Mother's education low	-0.032 (0.122)	-0.159 (0.149)	-0.132 (0.133)	-0.189 (0.167)
Mother's education medium	0.051 (0.067)	0.113 (0.067)	0.121 (0.067)	0.008 (0.070)
Not proceeding to University	-0.065 (0.097)	-0.048 (0.103)	-0.198** (0.076)	-0.178* (0.078)
Next-to-last year students	0.115 (0.096)	0.106 (0.103)	0.090 (0.105)	-0.009 (0.101)
Probability of completion	-0.007** (0.002)	-0.007** (0.002)	-0.001 (0.002)	-0.001 (0.002)
Constant	7.025*** (0.440)	7.214*** (0.460)	7.296*** (0.432)	8.112*** (0.431)
R2	0.055	0.078	0.042	0.061
N	578	568	613	592

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%. Reference group: Male, Dutch, father's education high, mother's education high, university.

Table 12. Variance coefficient regression VMBO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	0.109* (0.047)	0.022 (0.057)	0.154*** (0.036)	0.154*** (0.036)
Age	-0.003 (0.036)	-0.021 (0.042)	0.001 (0.024)	0.001 (0.024)
Immigrant	-0.056 (0.060)	-0.076 (0.070)	-0.044 (0.049)	-0.044 (0.049)
Father's education medium	-0.074 (0.058)	-0.081 (0.060)	-0.080 (0.043)	-0.080 (0.043)
Father's education low	-0.007 (0.082)	-0.021 (0.085)	-0.077 (0.059)	-0.077 (0.059)
Mother's education medium	0.076 (0.057)	0.051 (0.059)	0.058 (0.041)	0.058 (0.041)
Mother's education low	-0.081 (0.098)	-0.019 (0.103)	-0.056 (0.079)	-0.056 (0.079)
Continue: MBO	-0.042 (0.109)	-0.154 (0.111)	-0.156* (0.067)	-0.156* (0.067)
Continue: HBO	0.018 (0.114)	-0.057 (0.156)	-0.193** (0.074)	-0.193** (0.074)
Probability of completion	-0.001 (0.002)	0.000 (0.002)	-0.000 (0.001)	-0.000 (0.001)
Constant	0.801** (0.297)	0.932** (0.326)	0.739*** (0.165)	0.739*** (0.165)
R2	0.036	0.017	0.083	0.083
N	283	283	319	319

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%.
Reference group: Male, Dutch, father's education high, mother's education high, not continuing to tertiary education.

Table 13. Variance coefficient regression HAVO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.017 (0.072)	0.066 (0.044)	0.047* (0.023)	0.013 (0.026)
Age	-0.045 (0.035)	-0.027 (0.018)	-0.026 (0.014)	-0.031* (0.015)
Immigrant	-0.031 (0.044)	-0.053 (0.053)	0.018 (0.036)	0.023 (0.043)
Father's education medium	-0.056 (0.074)	0.017 (0.107)	0.034 (0.028)	0.021 (0.032)
Father's education low	-0.029 (0.070)	-0.079 (0.077)	0.032 (0.056)	0.025 (0.062)
Mother's education medium	0.011 (0.088)	0.127 (0.097)	-0.033 (0.026)	0.008 (0.029)
Mother's education low	-0.040 (0.083)	0.026 (0.083)	-0.123* (0.057)	-0.015 (0.075)
Continue: HBO	0.051 (0.071)	0.082 (0.087)	-0.010 (0.059)	0.033 (0.064)
Continue: University	0.038 (0.091)	0.104 (0.099)	0.031 (0.067)	0.093 (0.075)
Penultimate year students	0.012 (0.041)	-0.026 (0.063)	0.040 (0.025)	0.013 (0.028)
Probability of completion	-0.002 (0.002)	0.002 (0.002)	0.000 (0.001)	0.000 (0.001)
Constant	1.136** (0.395)	0.511* (0.254)	0.704*** (0.132)	0.716*** (0.148)
R2	0.010	0.011	0.034	0.019
N	660	660	729	705

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%.
Reference group: Male, Dutch, father's education high, mother's education high,
MBO or not continuing to tertiary education.

Table 14. Variance coefficient regression VWO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	0.060* (0.028)	0.060* (0.026)	0.059* (0.023)	0.029 (0.024)
Age	-0.037* (0.018)	-0.041* (0.019)	-0.035* (0.017)	-0.051** (0.017)
Immigrant	0.066 (0.046)	0.086 (0.050)	0.088* (0.038)	0.046 (0.042)
Father's education medium	0.062 (0.038)	0.031 (0.035)	0.025 (0.031)	0.056 (0.037)
Father's education low	0.003 (0.095)	0.055 (0.084)	-0.113 (0.062)	-0.095 (0.071)
Mother's education medium	-0.039 (0.032)	-0.036 (0.032)	-0.028 (0.028)	-0.022 (0.031)
Mother's education low	-0.037 (0.088)	-0.017 (0.091)	0.044 (0.060)	0.123 (0.068)
Not University	0.010 (0.032)	0.020 (0.032)	0.015 (0.027)	0.007 (0.029)
Penultimate year students	0.059 (0.033)	0.024 (0.033)	0.022 (0.028)	-0.004 (0.031)
Probability of completion	-0.001 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Constant	0.774*** (0.163)	0.761*** (0.159)	0.830*** (0.140)	0.976*** (0.151)
R2	0.050	0.039	0.046	0.044
N	569	569	608	595

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%.
Reference group: Male, Dutch, father's education high, mother's education high.

Table 15. Skewness coefficient regression VMBO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.067 (0.046)	-0.056 (0.053)	-0.057* (0.025)	-0.057* (0.025)
Age	-0.001 (0.026)	0.032 (0.035)	0.019 (0.017)	0.019 (0.017)
Immigrant	-0.064 (0.063)	0.025 (0.070)	0.025 (0.034)	0.025 (0.034)
Father's education medium	-0.051 (0.049)	0.048 (0.054)	0.040 (0.033)	0.040 (0.033)
Father's education low	-0.011 (0.081)	0.063 (0.093)	0.031 (0.048)	0.031 (0.048)
Mother's education medium	0.026 (0.050)	-0.037 (0.054)	-0.048 (0.032)	-0.048 (0.032)
Mother's education low	0.003 (0.073)	-0.036 (0.080)	-0.052 (0.053)	-0.052 (0.053)
Continue: MBO	-0.118 (0.155)	-0.034 (0.097)	-0.029 (0.069)	-0.029 (0.069)
Continue: HBO	-0.103 (0.163)	-0.109 (0.137)	0.018 (0.072)	0.018 (0.072)
Probability of completion	0.002 (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Constant	0.041 (0.236)	-0.007 (0.259)	0.015 (0.129)	0.015 (0.129)
R2	0.027	0.017	0.048	0.048
N	283	283	319	319

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%. Reference group: Male, Dutch, father's education high, mother's education high, not continuing to tertiary education.

Table 16. Skewness coefficient regression HAVO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	0.057 (0.076)	-0.041 (0.032)	-0.030* (0.015)	-0.030 (0.023)
Age	0.014 (0.037)	0.024 (0.015)	-0.004 (0.009)	0.008 (0.014)
Immigrant	0.024 (0.040)	0.025 (0.037)	-0.008 (0.022)	-0.018 (0.029)
Father's education medium	-0.029 (0.069)	-0.038 (0.067)	-0.010 (0.018)	0.002 (0.025)
Father's education low	-0.070 (0.074)	-0.006 (0.053)	-0.013 (0.029)	-0.022 (0.050)
Mother's education medium	0.127 (0.080)	-0.069 (0.042)	0.005 (0.017)	0.007 (0.026)
Mother's education low	0.121 (0.075)	0.109 (0.058)	-0.012 (0.041)	0.135** (0.049)
Continue: HBO	-0.074 (0.063)	-0.004 (0.062)	0.088** (0.034)	0.069 (0.049)
Continue: University	-0.078 (0.087)	-0.014 (0.068)	0.100* (0.042)	0.083 (0.063)
Penultimate year students	-0.037 (0.040)	0.051 (0.044)	-0.040* (0.017)	-0.005 (0.023)
Probability of completion	0.002 (0.002)	-0.001 (0.001)	0.000 (0.000)	-0.000 (0.001)
Constant	-0.196 (0.417)	0.083 (0.190)	0.009 (0.086)	0.060 (0.126)
R2	0.013	0.013	0.029	0.015
N	660	660	729	705

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%.
Reference group: Male, Dutch, father's education high, mother's education high, not continuing to tertiary education.

Table 17. Skewness coefficient regression VWO students

	Entry wage secondary education (I)	10 years after secondary education (II)	Entry wage tertiary education (III)	10 years after tertiary education (IV)
Female	-0.043 (0.027)	0.023 (0.023)	-0.048** (0.017)	-0.038 (0.022)
Age	0.016 (0.021)	0.008 (0.016)	0.005 (0.011)	0.007 (0.014)
Immigrant	-0.030 (0.053)	0.009 (0.041)	0.025 (0.024)	-0.002 (0.033)
Father's education medium	0.008 (0.036)	0.005 (0.029)	-0.011 (0.020)	0.005 (0.029)
Father's education low	-0.026 (0.080)	-0.023 (0.090)	-0.039 (0.062)	-0.216** (0.081)
Mother's education medium	0.061* (0.030)	0.015 (0.025)	-0.003 (0.018)	0.014 (0.025)
Mother's education low	0.069 (0.081)	-0.189** (0.073)	0.075 (0.053)	0.034 (0.055)
Not University	-0.021 (0.031)	-0.010 (0.024)	-0.016 (0.018)	0.043 (0.025)
Penultimate year students	0.025 (0.034)	-0.037 (0.026)	0.016 (0.018)	0.053* (0.024)
Probability of completion	0.001 (0.001)	-0.000 (0.001)	0.001 (0.001)	0.003*** (0.001)
Constant	-0.140 (0.181)	0.090 (0.149)	-0.023 (0.091)	-0.219 (0.136)
R2	0.019	0.025	0.030	0.061
N	569	569	608	595

Note: robust standard errors in parenthesis. **/** indicates significance levels at 10%/5%/1%.

Reference group: Male, Dutch, father's education high, mother's education high, university last year students.

Table 18. Descriptive statistics perceived returns to education by future destination

	N	Mean	S.D.	Min	Quantiles			
					.25	.50	.75	Max
Females								
Work	6	0.23	0.21	0.01	0.06	0.17	0.50	0.50
MBO	154	0.16	0.23	-0.60	0.02	0.13	0.27	1.00
HBO	507	0.13	0.17	-0.85	0.05	0.12	0.20	0.99
University	343	0.23	0.19	-0.54	0.12	0.21	0.32	1.00
Total	1010	0.17	0.19	-0.85	0.07	0.15	0.25	1.00
Males								
Work	4	0.21	0.21	0.06	0.08	0.12	0.33	0.52
MBO	74	0.19	0.23	-0.58	0.08	0.16	0.29	0.94
HBO	192	0.13	0.15	-0.68	0.05	0.13	0.22	0.67
University	155	0.21	0.17	-0.18	0.10	0.20	0.29	0.94
Total	425	0.17	0.18	-0.68	0.07	0.16	0.25	0.94

Table 19. Perceived returns to education

	VMBO	HAVO	VWO
	I	II	III
Female	-0.034 (0.030)	-0.004 (0.014)	0.013 (0.015)
Age	-0.012 (0.017)	0.008 (0.010)	-0.013 (0.011)
Immigrant	0.022 (0.046)	-0.019 (0.026)	-0.027 (0.021)
Father's education low	-0.050 (0.046)	-0.044 (0.025)	0.010 (0.061)
Father's education medium	-0.005 (0.042)	-0.025 (0.022)	-0.048* (0.020)
Mother's education low	0.114 (0.070)	-0.001 (0.047)	-0.001 (0.048)
Mother's education medium	-0.013 (0.041)	-0.007 (0.021)	0.001 (0.020)
Probability of completion	0.002** (0.001)	0.001 (0.001)	0.002* (0.001)
HBO	-0.125 (0.084)	-0.037 (0.074)	
MBO	-0.027 (0.081)		
University		-0.023 (0.077)	
Not proceeding to University			-0.110*** (0.014)
Penultimate year students		-0.026 (0.018)	-0.040* (0.018)
Constant	0.097 (0.105)	0.091 (0.083)	0.194* (0.096)
R2	0.081	0.028	0.099
N	260	616	549

Note: robust standard errors in parenthesis. */**/** indicates significance levels at 10%/5%/1%.
Reference group: Male, Dutch, father's education high, mother's education high, last year students.

Figure 1: Distribution of expected entry wages: males

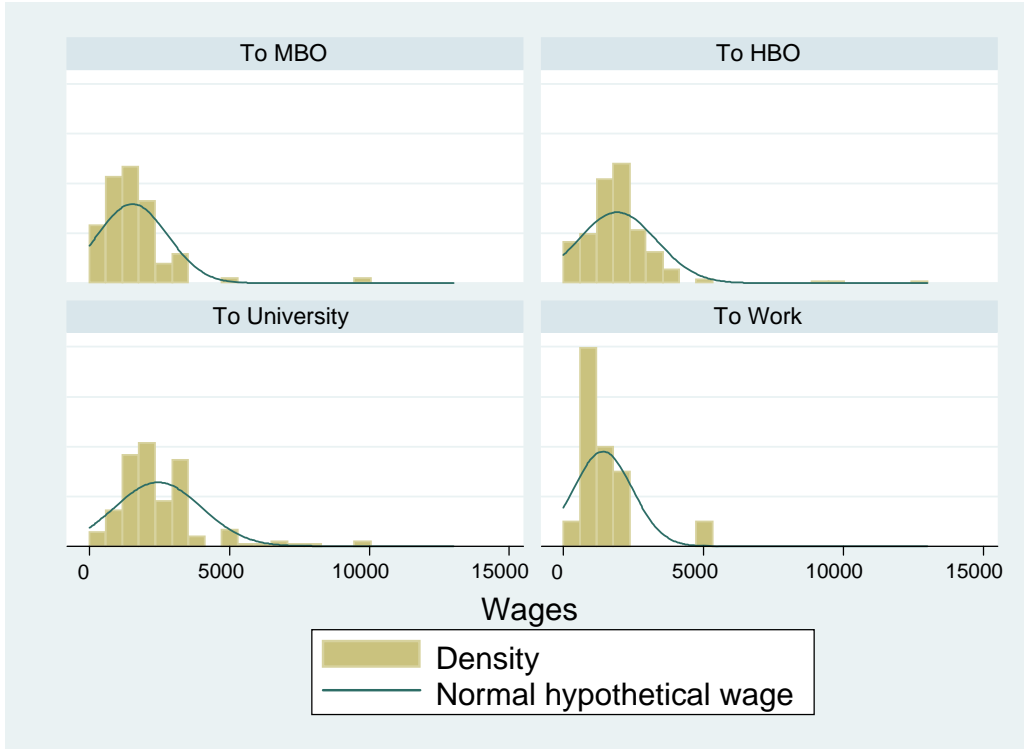
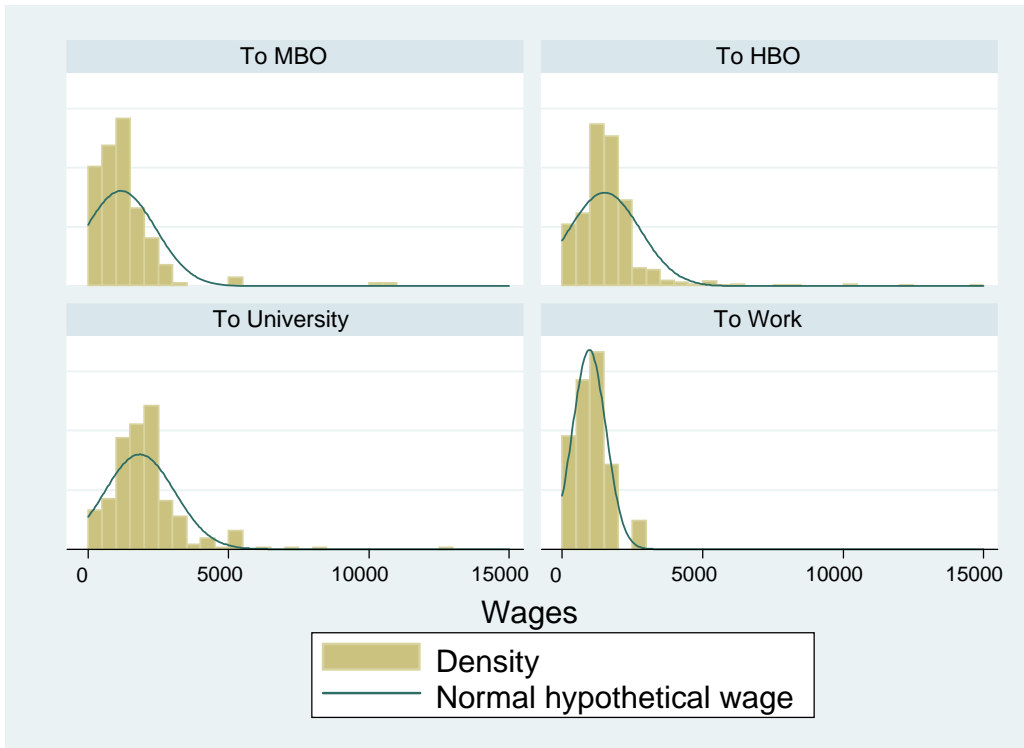


Figure 2: Distribution of expected entry wages: females



Note: we eliminated 5 observations for which the expected wage was equal to or exceeded 15000 Euros per month. The inclusion of those outliers would have caused the clustering of most of our observations in the left tail of the distribution.

Figure 3. Distribution of expected wages 10 years after entry: males

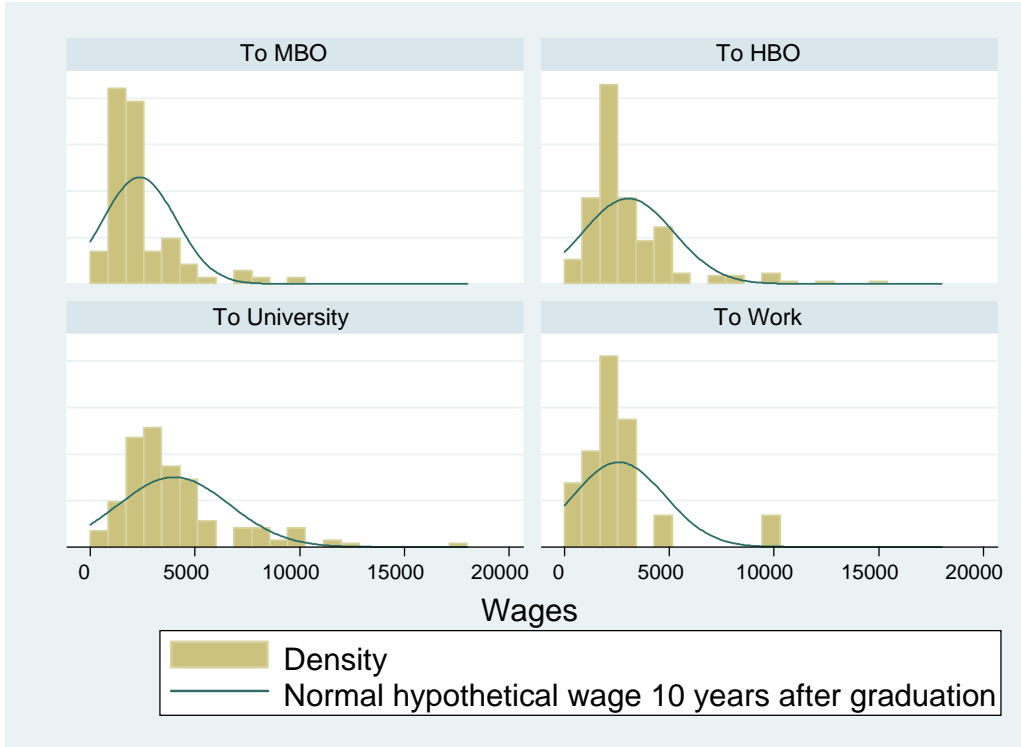
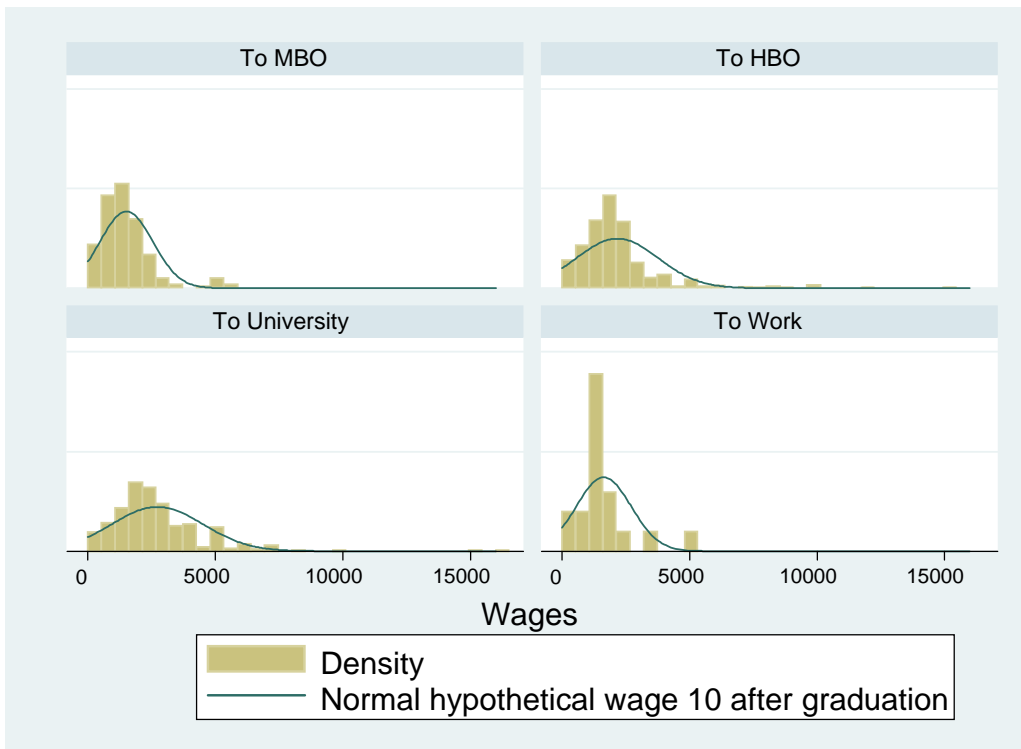


Figure 4. Distribution of expected wages 10 years after entry: females



Note: we eliminated 75 observations for which the expected wage was equal to or exceeded 20000 Euros per month. The inclusion of those outliers would have caused the clustering of most of our observations in the left tail of the distribution.

Figure 5

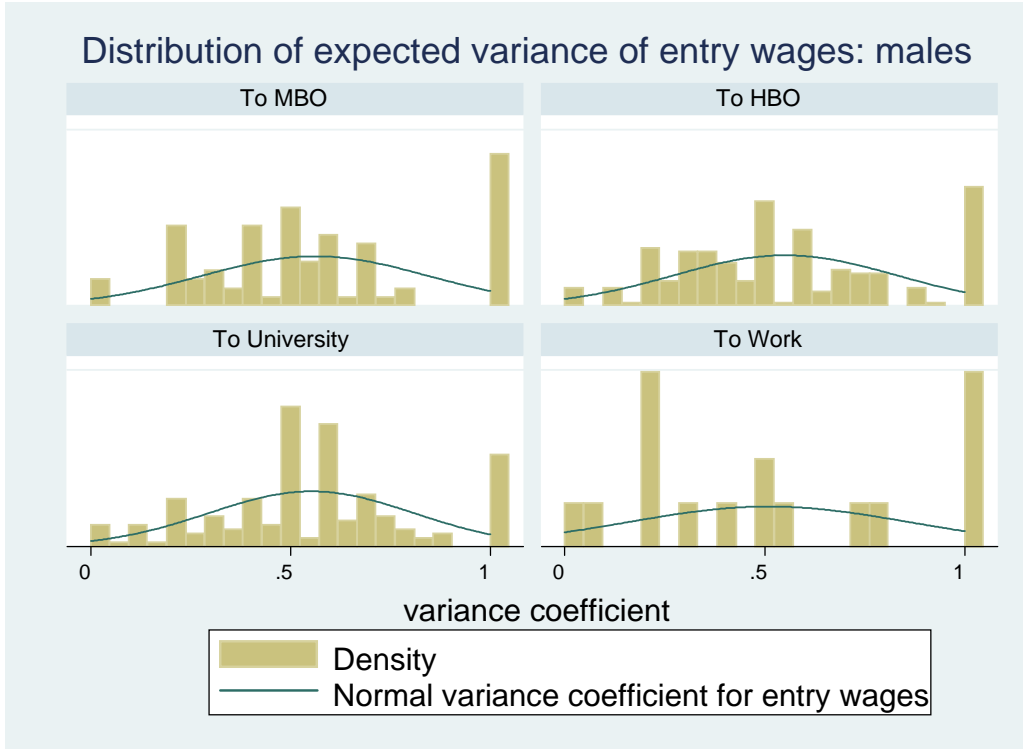


Figure 6

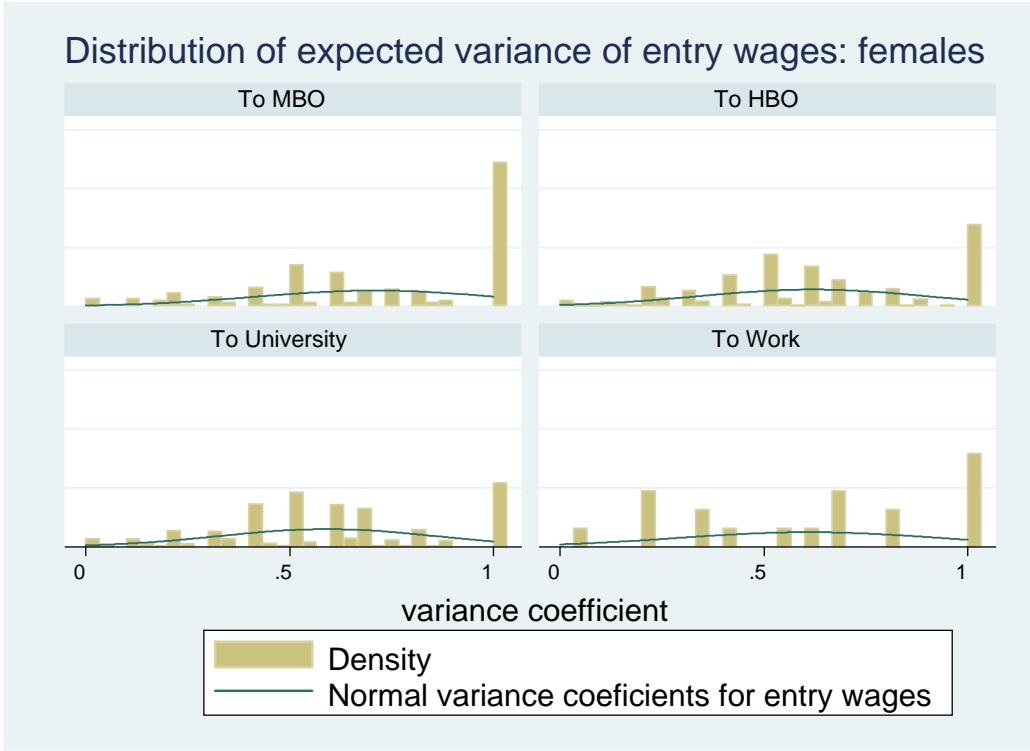


Figure 7

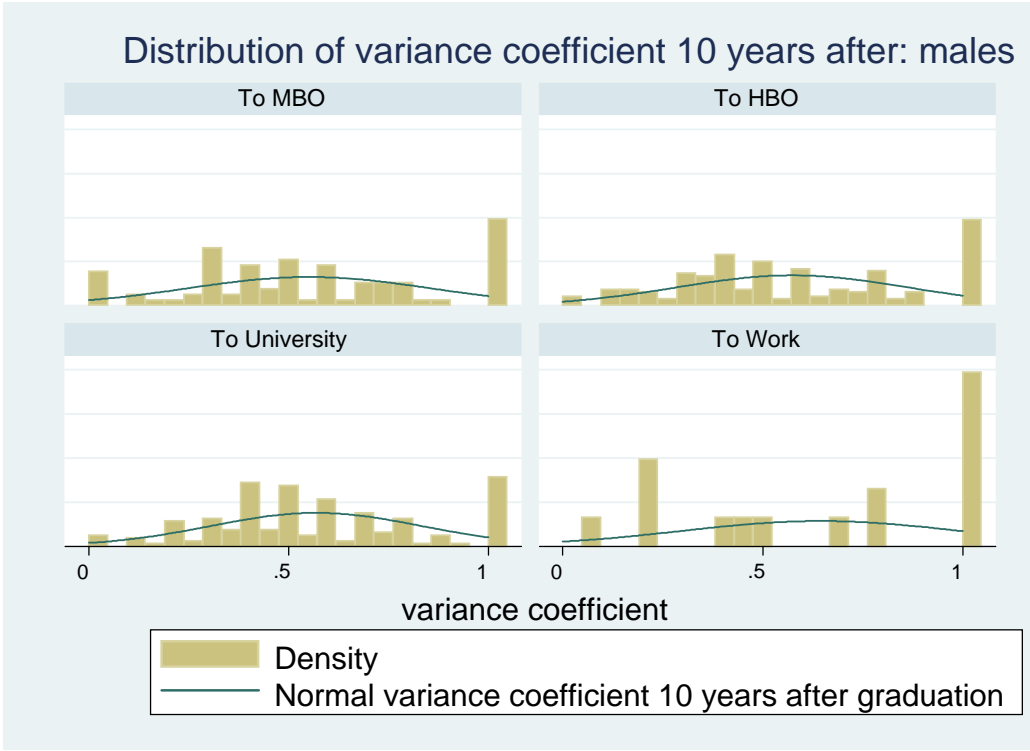


Figure 8

Distribution of variance coefficient 10 years after entry: females

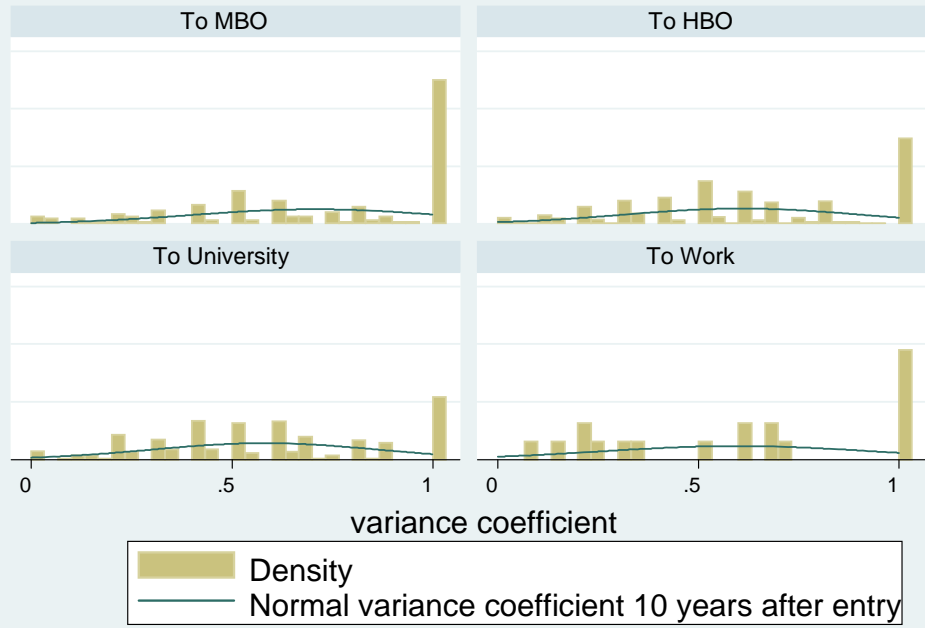


Figure 9

Distribution of skewness coeff. of entry wages: males

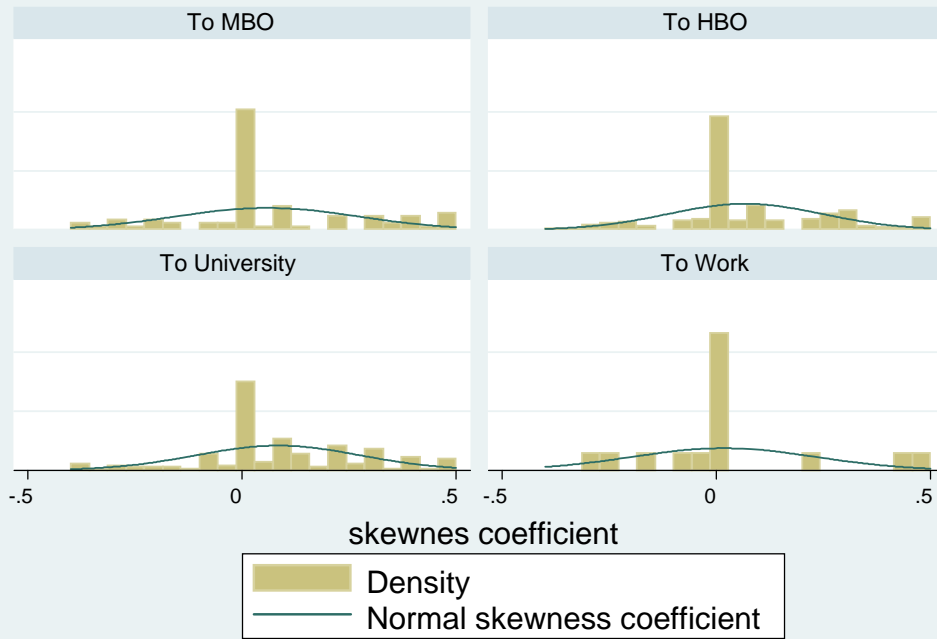


Figure 10

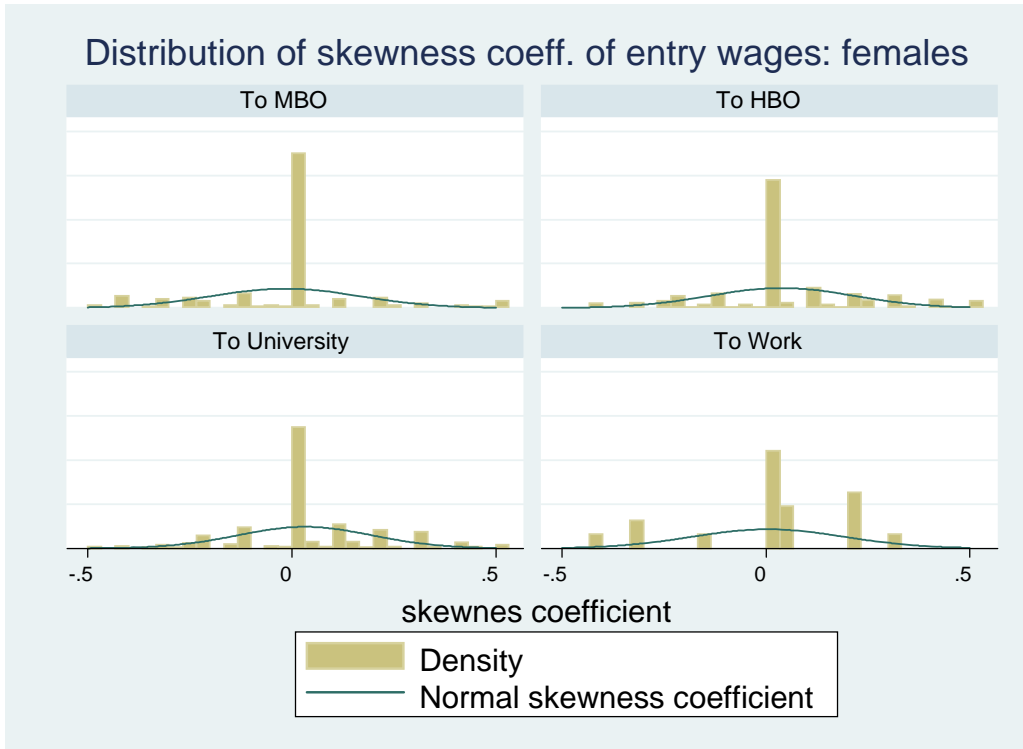


Figure 11

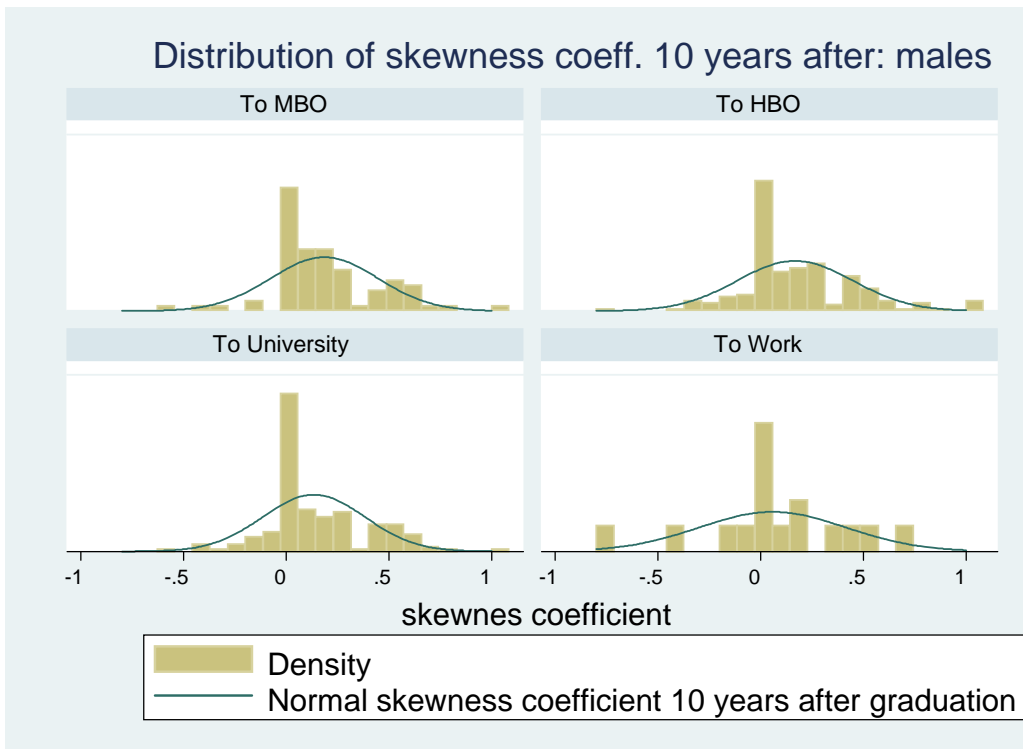


Figure 12

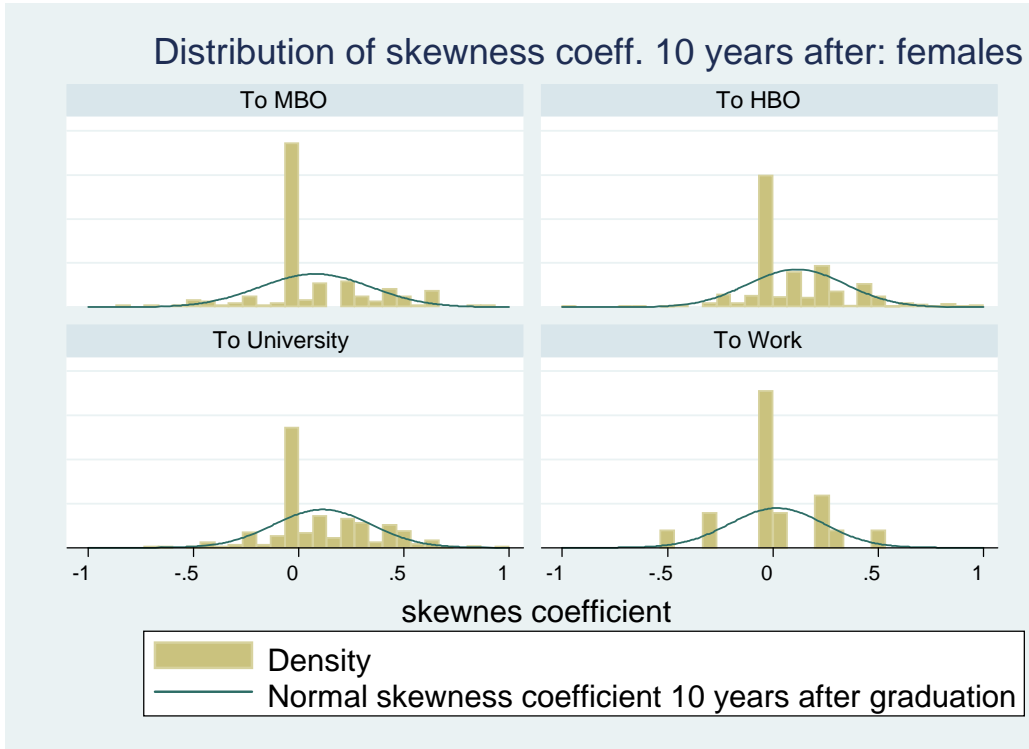


Figure 13

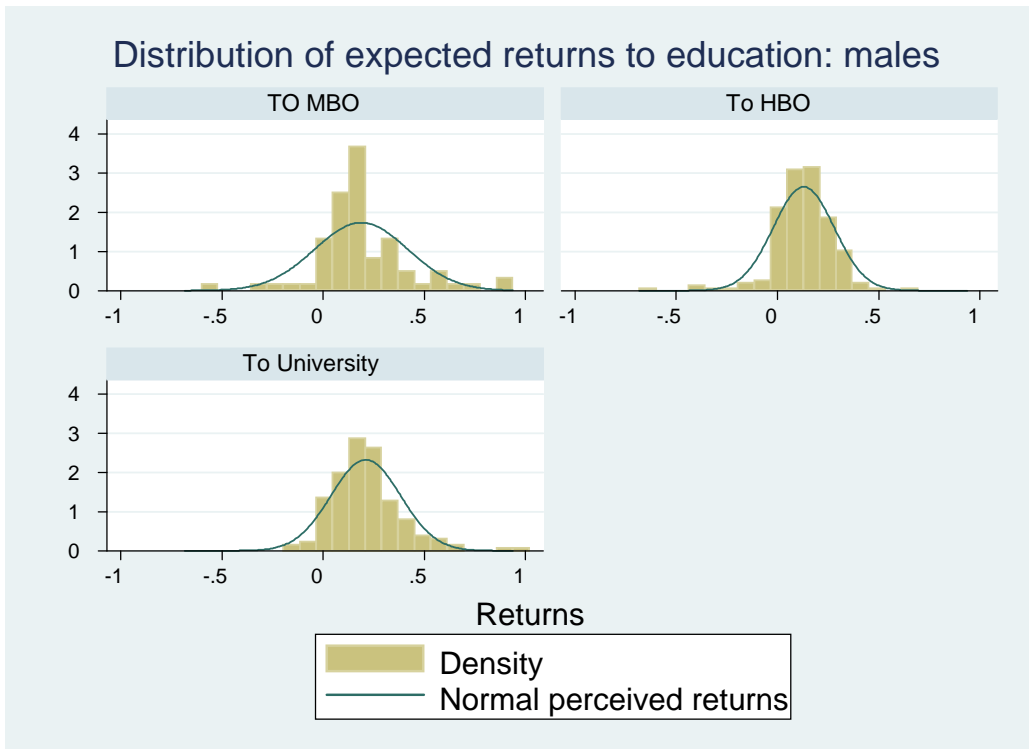
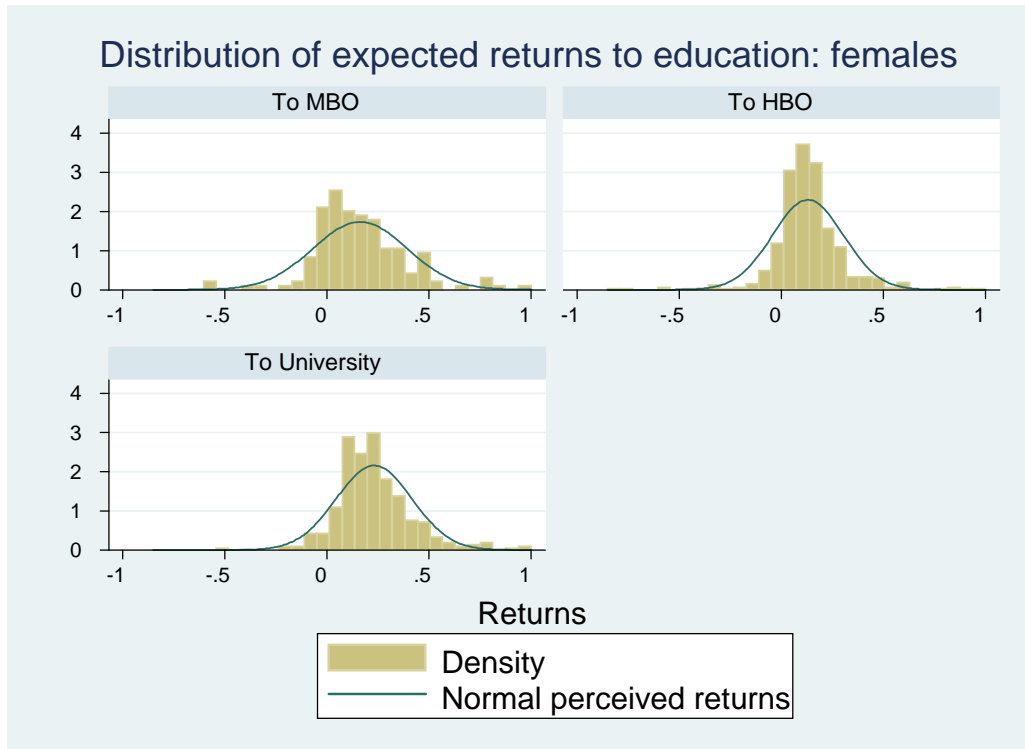


Figure 14



Appendix



Source: Wikipedia