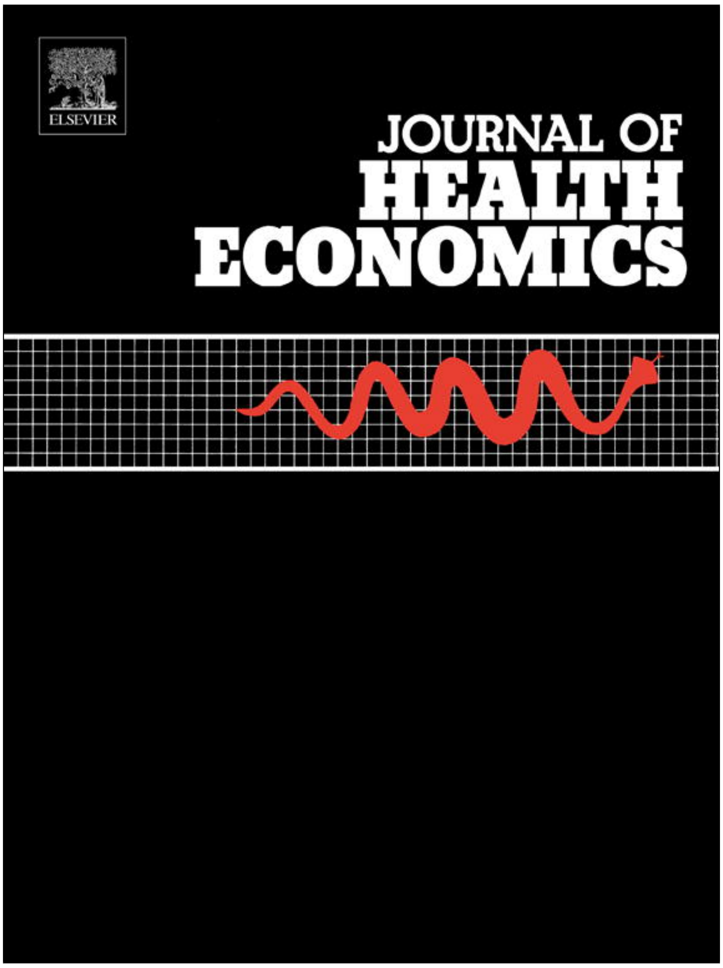
ELSEVIER
JOURNAL OF
HEALTH
ECONOMICS

Journal of Health Economics 32 (2013) 698–707

Contents lists available at SciVerse ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

# The effects of medical school on health outcomes: Evidence from admission lotteries

Edwin Leuven[a,1], Hessel Oosterbeek[b,*,2], Inge de Wolf[c]

[a] University of Oslo, Norway
[b] University of Amsterdam, The Netherlands
[c] Dutch Inspectorate of Education, The Netherlands

ARTICLE INFO

*Article history:*
Received 11 May 2012
Received in revised form 31 March 2013
Accepted 9 April 2013
Available online xxx

*JEL classification:*
I10
I23
I24

*Keywords:*
Education gradient
Admission lotteries
Medical school
Health outcomes
Organ donation

ABSTRACT

This paper estimates the effects of attending medical school on health outcomes by exploiting that admission to medical school in the Netherlands is determined by a lottery. Among the applicants for medical school, people who attended medical school have on average 1.5 more years of completed education than people who did not. They are also more likely to have been exposed to a health-related education curriculum. The results show only modest impacts on health outcomes. Attending medical school reduces alcohol consumption and being underweight somewhat, and has a small positive impact on self-reported health status. It has, however, a small negative effect on the frequency of physical exercise and no significant impact on smoking, and being overweight or obese. Attending medical school does have a large positive impact on the probability of being registered for donations of organs.



## 1. Introduction

This paper estimates the effects of attending medical school on health outcomes. For identification we exploit that admission to medical school in the Netherlands is determined by a lottery. People who won the lottery and attended medical school complete on average 1.5 more years of education than people who lost the lottery and attended some other study. They are also more likely to have been exposed to a health-related education curriculum. We are thus estimating the combined effect on health outcomes of 1.5 more years of post-secondary education and a more health-related curriculum.

Previous studies on the impact of education on health outcomes typically look at the impact of extra education at a lower level, for example by exploiting exogenous variation in years of schooling due to changes in compulsory schooling laws, Lleras-Muney (2005) for the US, Arendt (2005) for Denmark, Oreopoulos (2007) for the US, the UK and Canada, Albouy and Lequien (2009) for France, and Kemptner et al. (2011) for Germany. Park and Kang (2008) use variation in high school availability and birth order to identify a causal link of education on exercising and getting health checkups among Korean men. Most studies find a positive impact on health outcomes of extra education at this level.[3] While it is acknowledged that the content of education might matter for health, there is no causal evidence documenting this (Cutler and Lleras-Muney, 2008).

We use data from persons who applied for a medical study for the first time in the years 1988–1993 and who responded to a survey that was sent out in 2007. This means that most respondents are between 32 and 38 years old when information on health outcomes was collected. Since lottery losers can reapply in subsequent years and since some lottery winners choose not to

* Corresponding author. Tel.: +31 205254242.
*E-mail address:* h.oosterbeek@uva.nl (H. Oosterbeek).

[1] Leuven is also affiliated with CEPR, CESifo, IZA and Statistics Norway.

[2] Oosterbeek is also affiliated with TIER, Tinbergen Institute, CESifo and FLACSO-Ecuador.

[3] Webbink et al. (2010) use variation in the schooling levels of twins to establish an impact of education (average over different levels of education) on overweight and obesity of Australian men (but not women). Studies failing to find an impact of education on health include Tenn et al. (2010) and Braakmann (2011). See Grossman (2006) and Cutler and Lleras-Muney (2008) for reviews of the literature, including early studies that used less convincing identification strategies.

0167-6296/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jhealeco.2013.04.001

enroll in medical school, the correlation between the lottery result and medical school attendance is not perfect. We therefore use the result of the first lottery in which someone participated as an instrumental variable for medical school attendance.

Winning the first lottery increases the probability of attending medical school by 47 percentage points. Attending medical school, in turn, increases the length of formal education by 1.5 years, and increases the probability to enroll in a health-related program by 72 percentage points. The results show, however, only modest impacts on health outcomes. Attending medical school reduces alcohol consumption and being underweight somewhat, and has a small positive impact on self-reported health status. It has, however, a small negative effect on the frequency of physical exercise and no significant impact on smoking, and being overweight or obese. Finally, we also document that attending medical school has a substantial positive effect on individuals' altruistic health behavior, namely the probability that someone registers as a donor of organs.

Attending medical school may affect health outcomes directly through the length of education and the education content, but also indirectly through its effect on intermediate outcomes such as occupation, working hours, income, status and family situation. In the context of our study there are for example concerns about how well doctors care for themselves.[4] Like other studies that consider the effect of education on health, disentangling mechanisms is however not possible with one source of exogenous variation. We therefore have only limited possibilities to investigate the relative importance of the various channels.

To assess the possible role of different channels, we estimate the impact of attending medical school on variables that potentially mediate the influence of medical school on health outcomes. By and large, the results indicate that the differences in health outcomes cannot be due to the family situation; attending medical school has no impact on marital status and the number of children. Attending medical school does, however, have an impact on labor market outcomes. People who attended medical school work 8% more hours, but are not more likely to work more than 60 h per week. They also have 7.4% higher wages. Given that attending medical school has only a small effect on health outcomes, these results imply that the net effects of the mediating variables on health outcomes are also small. This may be due to mediating variables like work hours and wages, operating in opposite directions.

The remainder of this paper is organized as follows. The next section provides further details on the lottery for medical schools in the Netherlands and the institutional context. Section 3 describes the sources of data used in this paper. Section 4 outlines the estimation procedures and discusses identification issues. Section 5 presents and discusses the effects of attendance of medical school on health outcomes and behavior. It also discusses the evidence regarding possible channels. Section 6 summarizes and concludes.

## 2. The lottery and institutional context

University education in the Netherlands is provided by 13 universities. These universities are all publicly funded, offer programs of very similar contents and quality, and charge uniform tuition fees that are set by the government. Eight of these universities offer programs in medical education. Medical studies at the university level consist of a basic track of 4 years of pre-clinical training, followed by 2 years of clinical clerkships in hospitals. Graduates from this 6 years program get a first medical qualification (comparable with Doctor of Medicine) with which they can enter the labor market. After obtaining this degree students can also choose to continue their medical training. To become a general practitioner requires 3 extra years of training, whereas medical specializations like ophthalmology, radiotherapy or urology require 4–6 additional years. In order to get a place in one of the medical specialization tracks it is common to first get a PhD degree. In total, the complete medical study can take between 6 and 15 years. These specialization tracks are mainly on-the-job and those who partake in them have a work contract and receive a salary. About 80% of the specialization for general practitioner for example is on-the job.

Normally, all graduates from the pre-university track in secondary education can enroll in university in the field of their wish provided that their subject specialization in secondary school matches the chosen field of study. Only a limited number of university studies, medical studies being the most prominent one, have a fixed number of places available.[5] This leads to a shortage of places if the number of qualified applicants exceeds this fixed number. For all other fields of study, supply is supposed to accommodate demand.

With excess demand for a certain study program, available places are in most countries assigned through some form of selection based on merit. Instead of this, highly demanded seats in Dutch medical schools are assigned through a weighted lottery. While one may criticize this allocation system for ethical reasons or an alleged lack of efficiency, from a research perspective it has the advantage of creating a design that provides the opportunity to assess the effects of health education on various outcomes.[6]

Before the actual lottery takes place, applicants to medical studies in the Netherlands are assigned to lottery categories. The categories differ by the probability to be awarded a place (to win the lottery). For regular applicants with a Dutch pre-university diploma, six categories are distinguished. These categories are indicated by letters A to F and differ in the grade point average (GPA) applicants obtained for their final exams in secondary school. These exams are nation-wide and externally graded. Grades in Dutch secondary school are given on a scale from 1 to 10, where 6 and above indicate a pass. Non-passes in some subjects are allowed given sufficient compensation (above 6) on other subjects. The classification is as follows: A if GPA $\geq$ 8.5; B if 8.0 $\leq$ GPA < 8.5; C if 7.5 $\leq$ GPA < 8.0; D if 7.0 $\leq$ GPA < 7.5; E if 6.5 $\leq$ GPA < 7.0; F if 6 $\leq$ GPA < 6.5.

The ordering from A to F reflects differences in ability (probably including motivation). Because ability may have an independent effect on health outcomes, it is important that the analysis takes into account that assignment to medical school is only random conditional on lottery group.

Table 1 shows for each of the year cohorts included in our analysis, the proportions that have been admitted from each of the groups and the numbers of applicants per group. In each year around 85% of the applicants belongs to one of the groups D, E and F. Groups A and B are quite small. The probability to be admitted is close to 1 for applicants in category A, close to 0.90 for applicants in category B, and diminishes monotonically when going to C, D, E and F. In group F the odds are around 0.50. Hence, the probability to be admitted depends positively on GPA in the pre-university track. The year-to-year variation in fractions of admitted applicants (per

[4] Baldwin et al. (1997) for example argue that doctors do not take sick leave when ill and do not seek and receive proper medical treatment when needed. The authors supplement anecdotal accounts of this phenomenon with results from a longitudinal study of a class of young doctors, but since the study does not include a proper comparison group it is difficult to give a causal interpretation to the findings.

[5] Besides medical studies, fixed numbers of places are also present for dentistry, veterinary medicine and (in some years) technical business studies.

[6] Ketel et al. (2012) merge data from the admission lotteries to administrative income data to estimate the financial returns to medical school.

**Table 1**
Fraction admitted and number of applicants by year and lottery group, Cohorts 1989–1993.

| | | A | B | C | D | E | F | Total |
|---|---|---|---|---|---|---|---|---|
| 1988 | Pr(In) | 1.00 | 1.00 | 0.89 | 0.75 | 0.62 | 0.54 | 0.67 |
| | *In/Out* | *29/0* | *96/0* | *160/19* | *373/123* | *333/205* | *401/347* | *1393/694* |
| 1989 | Pr(In) | 1.00 | 1.00 | 0.96 | 0.80 | 0.66 | 0.58 | 0.71 |
| | *In/Out* | *30/0* | *84/0* | *151/7* | *344/86* | *349/182* | *405/292* | *1363/567* |
| 1990 | Pr(In) | 1.00 | 1.00 | 0.87 | 0.71 | 0.59 | 0.51 | 0.64 |
| | *In/Out* | *36/0* | *111/0* | *168/26* | *334/134* | *337/234* | *379/367* | *1365/761* |
| 1991 | Pr(In) | 1.00 | 0.88 | 0.76 | 0.63 | 0.50 | 0.43 | 0.55 |
| | *In/Out* | *41/0* | *115/15* | *153/48* | *342/205* | *326/323* | *361/500* | *1358/1091* |
| 1992 | Pr(In) | 1.00 | 0.84 | 0.71 | 0.59 | 0.48 | 0.42 | 0.52 |
| | *In/Out* | *51/0* | *95/18* | *168/67* | *353/247* | *330/359* | *432/604* | *1429/1295* |
| 1993 | Pr(In) | 0.93 | 0.72 | 0.62 | 0.51 | 0.41 | 0.36 | 0.45 |
| | *In/Out* | *41/3* | *120/47* | *150/91* | *355/347* | *346/501* | *466/833* | *1478/1822* |

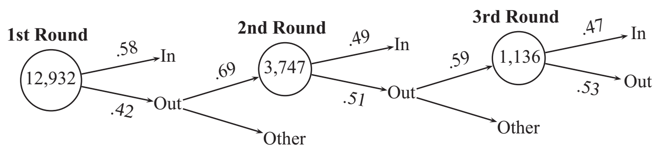


**Fig. 1.** Repeated participation, Cohorts 1989–1993

group) depends on the numbers of applicants and the total number of available slots. The small numbers of applicants in groups A and B in combination with the high fractions admitted cause that only very few applicants in these groups are not admitted (3 and 47 respectively in a period of 6 years). Because of this, these categories are excluded from our analyses.

Applicants who are not admitted in a given year, have the opportunity to reapply the next year. For the cohorts included in our analysis, there was no restriction on the total number of times someone could apply, and many indeed reapplied.[7] The repeated participation is illustrated in Fig. 1 which shows for 3 rounds lottery participation for the individuals in our dataset.[8] Of the 12,932 individuals we observe participating for the first time, 58% gets immediately admitted to medical school. Of the the 5430 applicants who do not get admitted in the first round, 69% reapplies the following year. We observe 3747 individual participating for the second time and the acceptance rate is now 49%. After an unsuccessful second round there is again a substantial fraction, 59%, that participates in the admission lotteries for a third time. Rejection rates go up because the pool of applicants consists increasingly of individuals from lower lottery categories. reapplication rates go down, probably because the opportunity costs of waiting one year to reapply increase. There are no possibilities to circumvent the lottery system and enroll in medical school through some other way.

As a consequence of repeated applications, admission to medical school is no longer completely random conditional on lottery category. Motivated applicants who are rejected the first time are more likely to reapply than applicants who are less motivated. We therefore make use of an instrumental variable approach where we use the lottery result of the first application as instrument for attending medical school (see Section 4 for details).

## 3. Data

The data used in this paper come from two sources. The first source is the administration of the agency that conducts the lottery. Starting in 1987, this source provides for each year, information about who applies for a medical study together with the lottery group and the outcome of the lottery. The same agency also registers enrollment in Dutch higher education, which informs us about the actual study choices of those who won the lottery as well as of those who lost the lottery. Information of study progress is available in the form of whether or not students successfully completed certain stages. From this we construct a variable that measures the number of years of completed education.

The second source is a survey that we designed and sent out in 2007 to all persons who applied for a medical study for the first time in the years 1988–1993. We start in 1988 rather than in 1987 because we want to make sure that we have administrative information about the results from the first lottery that someone participated in. For persons who applied in 1987, we cannot rule out that they applied but were not admitted in 1986.[9] In our analysis we further restrict our sample to applicants who are no older than 24. Because a medical study takes in total 6–15 years, we did not sample more recent cohorts, as many of these applicants would still be students.

A letter inviting people to participate in our research was sent to 10,475 applicants in groups C to F. It was communicated to respondents that responses to the survey would be merged to the administrative data. The response rate equals 0.56. We consider this a high response rate for a questionnaire that is not anonymous. The response rate is only slightly lower among applicants who were not immediately admitted than among applicants who were immediately admitted, 0.54 versus 0.57.[10]

The survey includes questions about health behavior and health status. Regarding health behavior, respondents were asked whether they smoked or not (and if so how much), how many alcoholic drinks they consume on average per week, and how frequently they do some physical exercise (sports). On the basis of this we created dummy variables for smoking and having

[7] Recently the number of (re-)applications has been restricted to three.

[8] This is based on the full sample, including applicants from groups A and B.

[9] The share of rejected applicants that skip one or more years before they apply again equals 2%. This indicates that only few of the 1987 applicants will be considered as first time applicants while in fact they are not.

[10] To assess whether response is selective, we regressed a dummy indicator for response on result of the first lottery, gender, age at first lottery, lottery category dummies and interaction terms. Table A1 in the appendix presents the estimation results. Column (1) shows that when we control for covariates, the difference in response rate between winners and losers of the first lottery is small (2 percentage points) but significant. Furthermore people who were older at the moment of application are somewhat less likely to respond while women are more likely to respond. The response rate is very similar across lottery groups but slightly higher for groups C and D. Column (2) shows that the coefficient of the interaction term between lottery result and lottery categories C and D is also significantly different from zero. The impact of the lottery result on the response rate is, however, never very large. All in all, these results show that response is non-random but that differences by observed variables are small, especially regarding the result of the first lottery.

**Table 2**
Descriptive statistics.

| | Scale | Attenders Mean | SD | Non-attenders Mean | SD | *p*-value |
|---|---|---|---|---|---|---|
| Age at interview | Years | 33.415 | 2.033 | 33.512 | 2.166 | 0.123 |
| Female | Dummy | 0.611 | 0.488 | 0.628 | 0.484 | 0.234 |
| Lottery-weight | 3 (highest)–6 (lowest) | 4.774 | 1.031 | 5.220 | 0.899 | 0.000 |
| *Health outcomes* | | | | | | |
| – Self-reported | 1 (very poor)–5 (excellent) | 4.318 | 0.634 | 4.270 | 0.614 | 0.009 |
| – BMI | Continuous | 23.104 | 3.048 | 23.236 | 3.136 | 0.152 |
| – Overweight (BMI > 25) | Dummy | 0.207 | 0.405 | 0.232 | 0.422 | 0.045 |
| – Obese (BMI > 30) | Dummy | 0.027 | 0.163 | 0.028 | 0.164 | 0.908 |
| – Underweight (BMI < 18) | Dummy | 0.011 | 0.103 | 0.012 | 0.107 | 0.791 |
| – Sport (# per week) | Categories | 1.279 | 1.102 | 1.394 | 1.112 | 0.001 |
| – Drinks per week | Continuous | 3.478 | 5.751 | 3.743 | 5.401 | 0.102 |
| – Drinks > 14 per week | Dummy | 0.044 | 0.206 | 0.050 | 0.218 | 0.362 |
| – Smoker | Dummy | 0.082 | 0.274 | 0.091 | 0.288 | 0.266 |
| – Number of cigarettes per day | Continuous | 9.148 | 6.915 | 10.289 | 7.468 | 0.117 |
| – Donor | Dummy | 0.697 | 0.460 | 0.561 | 0.496 | 0.000 |
| *Possible mediating variables* | | | | | | |
| – Single | Dummy | 0.155 | 0.362 | 0.174 | 0.379 | 0.079 |
| – Number of children | Continuous | 0.987 | 1.070 | 0.949 | 1.045 | 0.220 |
| – ln(working hours per week) | Continuous | 3.736 | 0.341 | 3.634 | 0.356 | 0.000 |
| – More than 60 h a week | Dummy | 0.103 | 0.304 | 0.078 | 0.268 | 0.004 |
| – ln(wages) | Continuous | 2.956 | 0.565 | 2.922 | 0.560 | 0.053 |
| – Nr of weeks unemployed | Continuous | 2.168 | 7.827 | 3.969 | 11.267 | 0.000 |
| – Years of completed education | Continuous | 4.620 | 1.740 | 3.328 | 1.692 | 0.000 |
| – Health-related study | Dummy | 1.000 | 0.000 | 0.295 | 0.456 | 0.000 |
| Number of observations | | 4286 | | 1556 | | |

more than 14 alcoholic drinks per week, continuous variables for the number of cigarettes per day and the number of alcoholic drinks per week, and an ordinal variable for frequency of physical exercise (number of times per week). From questions about respondents" weight and height we constructed their body mass index (BMI) and dummies for overweight (BMI > 25), obesity (BMI > 30) and underweight (BMI < 18) as indicators of health status.

To measure health status we asked respondents how they regard their own health situation in general on a 5 points scale: very poor, poor, neither good nor poor, good, excellent. This subjective information has been shown to be strongly correlated with objective health measures (Idler and Benyamini, 1997). Also related to health, we asked respondents whether they are officially registered as a donor of organs.[11] While this variable is not directly related to health behavior or health status but rather health altruism, we also analyze the effects of attending medical school on this variable. The hypothesis is that medical school attendance increases people's awareness of the importance of organ donations.

In search for mechanisms we also analyze variables that may potentially mediate the influence of medical school on health outcomes. The survey asked people about their family situation captured in a dummy for being single (as opposed to being married or cohabiting) and a continuous variable for the number of children. It also asks about the number of hours of work per week and earnings per month. We summarize the information in the logarithms of the number of hours and the hourly wage rate and a dummy for working more than 60 h per week. Furthermore the survey includes a question about unemployment spells, which we measure as the total number of weeks that a person has been unemployed.

Table 2 presents descriptive statistics for the health variables, control variables and possible mediating variables, separately for those who attended medical school and those who did not. A large share of the respondents report good or excellent health, leading to the high average level of health status in both groups. An important factor explaining this is the age of the respondents (33 on average with an SD of 2.4). According to the self-reported information on height and weight, 21% of the respondents are overweight, 1% underweight, and 3% obese. These percentages overweight and obese are low compared to fractions for the Dutch adult population in which 45% is overweight and 10% is obese. The median (modal) respondent has some physical exercise once (two or three times) per week. The average weekly intake of alcoholic beverages is 3–4 units; 4% of the respondents report to drink more than 14 units per week. Slightly more than 9% of the respondents report to smoke. In the Dutch population this percentage equals 28.

For some variables the responses are significantly different between respondents who attended medical school and those who did not. The most substantial but least surprising difference occurs for the weight category in the lottery for admission to medical school. Attenders have a significantly higher subjective health (although the size of the difference is minor), they sport significantly less often, a higher fraction is registered as organ donor, and if they smoke they smoke fewer cigarettes per day. There are also significant differences for some of the possible mediating variables, more specifically for the labor market variables and years of education completed. In the remainder of the paper we examine whether the differences in outcomes remain when covariates (including lottery weights) are included and when we correct for self-selection through reapplications.

Applicants who did not attend medical school end up in many different fields of study. 19% of the non-attending applicants attend a study in a professional college rather than at the university level; 3% of that group opts for the study most closely related to medical school, namely physiotherapy. 10% of the non-attending applicants decide not to study at all. Among the remaining 70% the most popular alternative university studies are psychology (9%) and law (6%).

[11] The Netherlands is among the countries where voluntary consent is determined by the "opt in" method.

## 4. Estimation

The estimation procedure applied in this paper is straightforward. We are interested in the effect ($\delta$) of attending medical school ($s_i$) on health ($y_i$). We postulate a linear relationship:

$$y_i = \delta s_i + x_i'\beta + \epsilon_i \tag{1}$$

where $\beta$ is a (vector of) parameters to be estimated, $\epsilon_i$ a disturbance term, and $x_i$ is a vector of control variables including an intercept, dummies for birth year, gender, lottery weight, year of first lottery. Because the lotteries are random conditional on lottery category we also include full interactions of lottery category and year of first lottery.[12]

Table A2

Estimation of $\delta$ using OLS with a random sample from the population (of higher educated people) may give a biased estimate of $\delta$ if medical school attendance ($s_i$) correlates with unobserved determinants of health outcomes ($\epsilon_i$). Endogeneity bias arises for instance when students who are more concerned about their health are also more inclined to attend medical school. Although in our sample this problem is probably greatly reduced by restricting the sample to persons who applied for medical school and by virtue of the admission lotteries, estimation of $\delta$ by OLS may still give biased estimates. As mentioned above, the complicating factor is that applicants who were rejected the first time can reapply the following year. Some of the rejected applicants decide to do so, while others do not to. As a consequence applicants who are admitted after their second application are no longer a random sample of the group of all applicants, and if the decision to reapply correlates with $\epsilon_i$ we might still end up with omitted variable bias.

To address this potential bias, we use the result of the first lottery ($z_i$) as instrumental variable for attending medical school. The result of the first lottery is a dummy that takes the value one for winners and equals zero for losers. The identifying assumption is that conditional on $x_i$, the result of the first lottery is mean independent of $\epsilon$ which is clearly fulfilled in case of the first lottery since at this point selective reapplication has not taken place yet.

The approach adopted in this paper allows us to identify a local average treatment effect (LATE, Imbens and Angrist, 1994). If the effect on applicants is homogeneous, it is equal to the average treatment effect on the treated (ATT). In that case we are able to answer the question what the health outcomes of those who attended medical school would have been if they would not have attended medical school. If treatment effects are heterogeneous, the effect that we estimate is the average treatment effect on the compliers, which is the group of people that attend medical school when they win the first lottery and do not attend medical school if they lose the first lottery.[13]

The research design does not inform us about the average effect (ATE) of attending medical school for a random sample from the entire population. The reason is that the population of applicants for medical school is not a random draw from the population.

In addition to instrumental variable estimates of Eq. (1) we also present naive OLS estimates of this relation as well as estimates of the first stage relation where attending medical school is the dependent variable and the result of the first lottery the explanatory variable of interest. Moreover, we present results from reduced form equations in which the health outcomes are the dependent variables and in which again the result of the first lottery is the explanatory variable of interest.

**Table 3**
Effects of lottery result on attending medical school (OLS).

| | Respondents to questionnaire | All applicants 1988–1993 |
|---|---|---|
| Result first lottery | 0.471*** | 0.487*** |
| | (0.011) | (0.008) |
| $R$-squared | 0.307 | 0.301 |
| $F$-test instrument | 1831.0*** | 3484.6*** |
| Number of observations | 5842 | 10,475 |

*Note*: Regressions include controls for year of birth, gender, lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. Observations in the two highest lottery categories (A and B) are excluded. Standard errors are heteroskedasticity robust.

*** Significance at the 1%-level.

As we discussed in the Introduction, the effect of attending medical school on health outcomes can operate through various channels. First of all, through the number of completed years of education and through the health content of the curriculum. The administrative data contain information about people's completed years of education which allows us to estimate the impact of medical school attendance on this variable. We also know whether people who lost the lottery enroll in another health-related study. This allows us to assess the impact medical school attendance on exposure to a study with a health-related curriculum.

As other mediating variables we consider:

- *Marital status*. Various studies claim a causal impact running from being married to health. In their survey of longitudinal evidence Wilson and Oswald (2005) suggest that the size of the health gain from marriage may be as large as the benefit from giving up smoking. See also the survey of Wood et al. (2007).
- *Number of children*. The presence of children may have an impact on adults' health status, although Ross et al. (1990) conclude in their survey that conditional on marital status, parents are not better off than non-parents.
- *Number of actual working hours per week and working more than 60 h per week*. While there is a longstanding research interest in the impact of working hours on health, Kirkcaldy et al. (2008) conclude that a clear-cut relationship has not emerged from that.
- *Income*. There is a high positive correlation between income and health, and various studies document that this is partly due to a causal impact of income on health (e.g. Frijters et al. (2005), Lindahl (2005)).
- Months without a job since graduation. There is a strong negative correlation between unemployment and health. A recent study by Schmitz (2011) fails, however, to find support for a causal impact of unemployment on health.

For each of these variables we will estimate the impact of attending medical school on them, using the result of the first lottery as instrumental variable. Whenever the impact of attending medical school on a variable equals zero, we can conclude that the channel associated with this variable cannot explain the relations that we find between attending medical school and health outcomes. Results are reported in Section 5.5.

[12] We also estimated models for each combination of lottery category and year of first lottery separately. Per health outcome, this results in 24 (4 categories times 6 years) different estimates. When we pool these estimates, by weighing with the number of observations, results are almost identical to the results reported in the main text but estimated with slightly less precision. See Table A2 in the appendix for details.

[13] The LATE interpretation assumes that the effect of the instrument on the endogenous variable is monotonic. This rules out that some individuals attend medical school when they lose the first lottery and don't attend medical school when they win the first lottery. This seems reasonable.

**Table 4**
Effects of lottery result and medical school on health outcomes (OLS, reduced form and 2SLS).

| | OLS (1) | OLS [POLS] (2) | RF (3) | 2SLS (4) | H-test (5) |
|---|---|---|---|---|---|
| *Health status* | | | | | |
| – Self-reported | 0.038** | 0.071*** | 0.015 | 0.033 | 0.876 |
| | (0.019) | (0.028) | (0.017) | (0.037) | |
| – Overweight | −0.009 | | −0.005 | −0.010 | 0.955 |
| | (0.012) | | (0.011) | (0.024) | |
| – Obese | 0.003 | | 0.002 | 0.005 | 0.880 |
| | (0.005) | | (0.004) | (0.010) | |
| – Underweight | −0.001 | | −0.005* | −0.011* | 0.076 |
| | (0.003) | | (0.003) | (0.007) | |
| *Health behavior* | | | | | |
| – Drinks | −0.298* | | −0.321** | −0.680** | 0.188 |
| | (0.159) | | (0.151) | (0.320) | |
| – Drinks > 14 | −0.007 | | −0.012** | −0.026** | 0.066 |
| | (0.006) | | (0.006) | (0.012) | |
| – Smoker | −0.002 | −0.205* | −0.012 | −0.025 | 0.100 |
| | (0.009) | (0.017) | (0.008) | (0.016) | |
| – # Cigarettes | −1.389* | | −0.052 | −0.104 | 0.271 |
| | (0.748) | | (0.673) | (1.341) | |
| – Sport | −0.109*** | −0.364*** | −0.049 | −0.104 | 0.928 |
| | (0.034) | (0.057) | (0.031) | (0.065) | |
| *Health altruism* | | | | | |
| – Donor | 0.120*** | | 0.053*** | 0.112*** | 0.759 |
| | (0.015) | | (0.013) | (0.028) | |

*Note*: Number of observations in columns (1), (3) and (4) equals 5842 or slightly less due to partial non-response. Number of observations in column (2) equals 13,869 or slightly less due to partial non-response. All coefficients come from separate regressions. The regressions include controls for year of birth, gender, and in columns (1) (3) and (4) for lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. Column (5) reports the *p*-value from a regression based Hausman test. Standard errors are heteroskedasticity robust.

* Significance at the 10%-level.
** Significance at the 5%-level.
*** Significance at the 1%-level.

## 5. Results

This section is divided into five subsections. We first present first stage results of the impact of the result of the first lottery on medical school attendance. We then discuss results from OLS estimation, followed by the reduced form estimates and the 2SLS result. We end this section with discussing possible channels.

### 5.1. First stage

Table 3 reports first stage results. The first column shows that for the respondents in our sample, winning the first lottery increases the probability to attend medical school by 47 percentage points. If all first lottery winners would attend medical school and if none of the first lottery losers would do so, this estimate would be equal to 1. Of the 3333 winners of the first lottery, 201 decided not to attend medical school. Of the 2509 persons drawing a zero in the first lottery, 1154 have been admitted after participating in lotteries of subsequent years. The second column uses information from all applicants in the period 1988–1993, including the people who did not respond to the survey. This shows that in terms of the first stage relationships, the respondents to the questionnaire are very similar to the population of applicants in this period.

### 5.2. OLS

Before we turn to the 2SLS estimates of the effect of attending medical school on health outcomes, we first discuss the naive OLS estimates. These results show us what we would conclude if we ignore that reapplications may give rise to selection bias. Column (1) of Table 4 reports the results from regressions of the health outcomes on a dummy variable for attending medical school and control variables. We see significant associations between attending medical school and self-reported health status (+), drinking (−), number of cigarettes in case of smoking (−), physical exercise (−) and being a donor of organs (+).[14]

The OLS estimates in Table 4 are already purged from selection due to differences in for example motivation, other intrinsic characteristics, and different subject specialization in secondary school between people who did apply to medical school and people who did not. To get an impression of the size of the bias that is taken away by this, we also estimated OLS regressions using higher educated people who did not apply to medical school as the comparison group. Data come from 5 waves of a Dutch survey known as POLS (Ongoing Survey on Living Conditions), which asks questions about some health related variables (self-reported health status, smoking and frequency of physical exercise). It contains information from a representative sample of the adult population. We restricted the sample to observations in the age range of 30–40 with a higher education degree, leaving us with an alternative comparison group of 6505 observations. We regressed the health variables on an indicator of attending medical school, including controls for gender and age.

Results from OLS-regressions using the alternative comparison group are reported in column (2) of Table 4. For self-reported health status the coefficient of attending medical school is substantial and highly significant. The difference amounts to more than 10% of a standard deviation of subjective health status. For smoking as the outcome variable, the coefficient suggests that attending medical school reduces smoking by 20 percentage points. For sporting the

[14] In OLS regressions of health outcomes on attendance of medical school and GPA (instead of dummies for lottery categories and their interactions with lottery years), the estimates for medical school are very similar to those in column (1) of Table 4. The coefficients of GPA are significant in regressions for overweight (−), obese (−), smoking (−) and donor (+).

**Table 5**
Gender differences (2SLS).

| | Men | | Women | | Difference | |
|---|---|---|---|---|---|---|
| | (1) | | (2) | | (1)–(2) | |
| *Health status* | | | | | | |
| – Self-reported | 0.039 | (0.060) | 0.033 | (0.047) | 0.006 | (0.076) |
| – Overweight | −0.064 | (0.044) | 0.024 | (0.028) | −0.089* | (0.052) |
| – Obese | 0.009 | (0.015) | 0.002 | (0.012) | 0.007 | (0.019) |
| – Underweight | −0.002 | (0.004) | −0.016 | (0.010) | 0.014 | (0.011) |
| *Health behavior* | | | | | | |
| – Drinks | −0.950 | (0.702) | −0.529* | (0.282) | −0.422 | (0.756) |
| – Drinks > 14 | −0.041 | (0.029) | −0.015 | (0.009) | −0.027 | (0.031) |
| – Smoker | −0.032 | (0.033) | −0.021 | (0.017) | −0.011 | (0.037) |
| – # Cigarettes | −1.543 | (1.860) | 1.696 | (1.983) | −3.239 | (2.719) |
| – Sport | −0.120 | (0.109) | −0.098 | (0.081) | −0.022 | (0.136) |
| *Health altruism* | | | | | | |
| – Donor | 0.173*** | (0.046) | 0.076** | (0.035) | 0.097* | (0.058) |

*Note*: Coefficients in the first two columns come from separate regressions. These regressions include controls for year of birth, lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. Standard errors are heteroskedasticity robust.
* Significance at the 10%-level.
** Significance at the 5%-level.
*** Significance at the 1%-level.

impact is negative and significant; people who attended medical school exercise on average 0.36 times less per week. Hence, for all three outcomes the estimates are much larger in absolute size when we use a representative sample of higher educated people of the same age as comparison group than when we compare with people who applied for medical school but didn't get in.

### 5.3. Reduced form estimates

This subsection discusses results from reduced form equations that estimate the effects of the outcome of the first lottery (the instrument) on health outcomes. These are the effects of the Dutch allocation system of places in medical school through a lottery on the health outcomes of winners of their first lottery. For all outcomes, we estimated linear models controlling for age and gender and for dummies of lottery categories, dummies for year of first lottery and their interactions. Results are reported in column (3) of Table 4. Results from (ordered) probit models are very similar.

Looking first at health status we find that for self-reported health status, overweight and underweight the effects of winning the first lottery have the expected signs. For obesity the sign is, however, in the unexpected (positive) direction. With the exception of being underweight, in all cases, however, the point estimates are (much) smaller than their standard errors. Although these effects are insignificant, they are small and if we add or subtract two standard errors to or from the point estimate in the direction of positive health effects of winning the lottery, we can still exclude substantial effects. With 95% probability, the effect of winning the lottery on self-reported health status is less than 7% of a standard deviation of the self-reported health status variable; and with the same probability the effects of winning the lottery on overweight and obesity are less than 3% and 11% of their respective standard deviations.

Winners are significantly less likely to drink (heavily) than lottery losers. For smoking and physical exercise no effects are found of winning the lottery. There is a substantial and statistically significant effect of winning the lottery on the probability of being registered as donor of organs. This probability is 5 percentage points higher, relative to a base of 62% which is the fraction of registered donors in the group that lost their first lottery.[15]

### 5.4. 2SLS

Due to possible endogeneity of attendance of medical school conditional on applying, the associations in column (1) can perhaps not be given a causal interpretation. The results in column (4), however, can be interpreted as such. The point estimates in this column are equal to those in column (3) divided by the first stage effects of winning the lottery on attendance in Table 3 (0.47). This leads to more than a doubling of the reduced form estimates. The standard errors in column (4) are also more than doubled in magnitude compared to those in column (3). Column (4) shows that those who attended medical school are less likely to be underweight and drink less than those who did not attend medical school. Those who attended medical school are also significantly more likely to be registered as a donor of organs. The effects size on drinking is a bit over 10% of its standard deviation. The effect on the probability to be registered as organ donor, equals more than 11 percentage points.[16]

Most 2SLS estimates in column (4) are close to the corresponding OLS estimates in column (1). This suggests that selective participation in later lotteries may not be a big concern, at least with respect to potential health outcomes. In the absence of endogeneity, OLS (and 2SLS) will estimate the average effect for the population of medical school applicants. We formally test for endogeneity using a regression based Hausman test, the p-values of which are reported in column (5). We reject the null-hypothesis of no endogeneity at the 10% level for two outcomes: underweight, and drinking more than 14 alcoholic beverages per week. For the remaining outcomes with similar points estimates in columns (1) and (4) we can thus consider the more precise OLS estimates. Apart from the fact that where we find effects we can interpret them as

[15] To assess the sensitivity of our findings for non-response we estimated the bounds proposed by Lee (2009) for the reduced form effects. Table A3 in the appendix reports the 95% confidence interval on the lower and upper bounds. These confidence intervals are somewhat wider than those of the reduced form estimates in column (3) of Table 4.

[16] To further assess the sensitivity of our findings for non-response, Table A4 in the appendix reports 2SLS results from a specification that excludes gender and age from the set of control variables. Although gender and age are strong predictors of response to the questionnaire (see Table A1), exclusion of these variables from the specifications gives virtually the same estimates. Column (3) in Table A4 reports results based on inverse probability weighted 2SLS, where the response probability is estimated using a Probit model with full interactions between lottery result, age, gender and lottery category. Again the results are almost identical. These results suggest that non-response bias is not a concern.

average effects for the pool of applicants, we now also find a small (6% of a standard deviation) but statistically significant effect of medical school attendance on health status, and a negative effect on sport.

Table 5 reports separate 2SLS-estimates for men and women. Gender differences in the health effects of attending medical school are never significantly different from zero. Notice, however, that for all outcome variables the effects estimates are larger (in absolute size) for men than for women. The difference is largest for the effect on being registered as organ donor; the effect for men is double that of women.

We also tested for differences by ability (not reported here), but did not find a systematic pattern between ability levels (measured by GPA in secondary school) and the size of the impact of health education on health outcomes. Only for being a smoker we saw a monotonic pattern with the effect of attending medical school being more beneficial for applicants of lower ability. The standard errors on the estimates were however too large for these differential effects to be statistically significant.

### 5.5. Mechanisms

Table 6 reports 2SLS estimates of the impact of attending medical school on variables that can potentially mediate an effect of attending medical school on health outcomes.

The first two rows show that those who attended medical school complete 1.5 more years of education that those who did not, and that those who attended medical school are 72% more likely to enroll in a health-related study. This confirms that attending medical school increases the amount of completed education and the exposure to a health-related curriculum.

We find no support for an effect through a change in the family situation; attending medical school has no impact on marital status and the number of children. Attending medical school does, however, have an impact on labor market outcomes. People who attended medical school work 8% more hours, but are not more likely to work more than 60 h per week. They also have 7.4% higher wages.

Given that we find rather small effects of attending medical school on health outcomes, these results imply that the net effects of the mediating variables on health outcomes are also small. This may be the result of opposing effects of years of education, a health-related curriculum and income in one direction and of working hours in the other direction. It may also be the result of each of these variables having just a minor effect on health outcomes. The latter might be true despite the fact that other studies find substantial effects of years of education, income and working hours on health. In this regard it is important that the levels of these variables amongst the people in the control group are already quite high (cf. Table 2). Our research design does not allow us to further explore the relative importance of the different channels.

**Table 6**
Effects of medical school on mediating variables (2SLS)

| Outcome | Medical school | |
|---|---|---|
| *Education* | | |
| – Years of completed schooling | 1.466*** | (0.084) |
| – Health-related study | 0.720*** | (0.015) |
| *Family situation* | | |
| – Single | −0.028 | (0.021) |
| – Nr of children | 0.043 | (0.059) |
| *Labor market situation* | | |
| – ln(hours worked) | 0.080*** | (0.019) |
| – 60+hour working week | 0.018 | (0.018) |
| – ln(hourly wage) | 0.074** | (0.033) |
| – Nr of weeks unemployed | −0.968* | (0.533) |

*Note*: Number of observations equals 5842 or slightly less due to partial non-response. All coefficients come from separate 2SLS-regressions. Medical school attendance is instrumented by the result of the first lottery. Regressions include controls for year of birth, gender, lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. Standard errors are heteroskedasticity robust.
* Significance at the 10%-level.
** Significance at the 5%-level.
*** Significance at the 1%-level.

## 6. Conclusion

The main contribution of this paper is to provide new evidence on the effects of education on health outcomes. Since the analysis exploits a lottery, the identification strategy is arguably more credible than most other studies on the overall effects of education on health. In addition, few of the existing studies are informative about the effect of post-secondary education, and this study adds to this. In particular we consider medical education, which is interesting because its content is directly relevant to the

**Table A1**
Effects of predetermined variables on survey response.

| | (1) | | (2) | |
|---|---|---|---|---|
| Result first lottery | 0.020** | (0.010) | 0.198 | (0.164) |
| Age at first lottery | −0.040*** | (0.004) | −0.035*** | (0.006) |
| Female | 0.079*** | (0.010) | 0.093*** | (0.015) |
| Lottery weight (F = reference) | | | | |
| C | 0.067 | (0.044) | 0.005 | (0.057) |
| D | 0.061* | (0.032) | 0.026 | (0.037) |
| E | −0.002 | (0.032) | −0.011 | (0.035) |
| Interaction terms | | | | |
| – Result first lottery × Age at first lottery | | | −0.010 | (0.008) |
| – Result first lottery × Female | | | −0.025 | (0.020) |
| – Result first lottery × Lottery weight C | | | 0.076* | (0.043) |
| – Result first lottery × Lottery weight D | | | 0.051* | (0.026) |
| – Result first lottery × Lottery weight E | | | 0.015 | (0.025) |
| Adjusted $R$ squared | 0.019 | | 0.020 | |
| Number of observations | 10,475 | | 10,475 | |

*Note*: Regressions include controls for interactions of lottery weight and year of first lottery. Observations in the two highest lottery categories (A and B) are excluded. Standard errors are heteroskedasticity robust.
* Significance at the 10%-level.
** Significance at the 5%-level.
*** Significance at the 1%-level.

**Table A2**
Effects of medical school on health outcomes; pooling observations/estimates from different lotteries.

| | Pooled observations | | Pooling estimates | |
|---|---|---|---|---|
| *Health status* | | | | |
| – Self-reported | 0.033 | (0.037) | 0.040 | (0.045) |
| *Health behavior* | | | | |
| – Overweight | −0.010 | (0.024) | −0.010 | (0.028) |
| – Obese | 0.005 | (0.010) | 0.004 | (0.010) |
| – Underweight | −0.011$^{*}$ | (0.007) | −0.010 | (0.007) |
| – Drinks | −0.680$^{**}$ | (0.320) | −0.525 | (0.367) |
| – Drinks > 14 | −0.026$^{**}$ | (0.012) | −0.033$^{**}$ | (0.015) |
| – Smoker | −0.025 | (0.016) | −0.023 | (0.018) |
| – # Cigarettes | −0.104 | (1.341) | | |
| – Sport | 0.122$^{*}$ | (0.064) | 0.126 | (0.079) |
| *Health altruism* | | | | |
| – Donor | 0.112$^{***}$ | (0.028) | 0.105$^{***}$ | (0.033) |

*Note*: All coefficients in column (1) come from separate 2SLS regressions which include controls for year of birth, gender, lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. The estimates in column (2) are the sample weighted averages of the estimates from 2SLS regressions stratified by lottery weight and year. Standard errors are heteroskedasticity robust. Outcomes in column (2) are also corrected for year of birth and gender.

$^{*}$ Significance at the 10%-level.
$^{**}$ Significance at the 5%-level.
$^{***}$ Significance at the 1%-level.

outcomes we consider. Attending medical school reduces alcohol consumption and being underweight somewhat, and has a small positive impact on self-reported health status. It has, however, a small negative effect on the frequency of physical exercise and no significant impact on smoking, and being overweight or obese. Attending medical school does have a large positive impact on the probability of being registered for donations of organs.

It has been claimed that young doctors are not taking sick leave when ill and do not seek or receive proper care when needed. Our results indicate that this is not due to working a doctor. Attending medical school has no negative impact on health outcomes, and if anything the health outcomes are slightly improved. It may be that the long working hours of doctors are harmful for their health outcomes but if that is the case, this is compensated by positive

**Table A3**
Assessing the importance of response bias using (Lee, 2009) bounds.

| | RF (1) | | RF (2) | | Selection corrected 95% CI | |
|---|---|---|---|---|---|---|
| *Health status* | | | | | | |
| – Self-reported | 0.015 | (0.017) | 0.017 | (0.017) | [−0.045 | 0.102] |
| *Health behavior* | | | | | | |
| – Overweight | −0.005 | (0.011) | −0.003 | (0.011) | [−0.059 | 0.024] |
| – Obese | 0.002 | (0.004) | 0.002 | (0.005) | [−0.028 | 0.011] |
| – Underweight | −0.005$^{*}$ | (0.003) | −0.005$^{*}$ | (0.003) | [−0.026 | 0.002] |
| – Drinks | −0.321$^{***}$ | (0.151) | −0.244 | (0.156) | [−1.093 | 0.162] |
| – Drinks > 14 | −0.012$^{**}$ | (0.006) | −0.010$^{*}$ | (0.006) | [−0.042 | 0.004] |
| – Smoker | −0.012 | (0.008) | −0.010 | (0.008) | [−0.056 | 0.005] |
| – # Cigarettes | −0.052 | (0.673) | −0.039 | (0.670) | | |
| – Sport | −0.049 | (0.031) | −0.048 | (0.031) | [−0.199 | 0.051] |
| *Health altruism* | | | | | | |
| – Donor | 0.053$^{***}$ | (0.013) | 0.052$^{***}$ | (0.013) | [0.020 | 0.106] |
| *Family situation* | | | | | | |
| – Single | −0.004 | (0.006) | −0.004 | (0.006) | [−0.073 | 0.005] |
| – Nr of children | 0.020 | (0.028) | 0.017 | (0.028) | [−0.119 | 0.123] |
| *Labor market situation* | | | | | | |
| – ln(hours worked) | 0.038$^{***}$ | (0.009) | 0.043$^{***}$ | (0.010) | [−0.006 | 0.109] |
| – 60+hour working week | 0.009 | (0.008) | 0.010 | (0.009) | [−0.046 | 0.030] |
| – ln(hourly wage) | 0.035$^{**}$ | (0.016) | 0.034$^{**}$ | (0.016) | [−0.059 | 0.219] |
| – Nr of weeks unemployed | −0.457$^{*}$ | (0.252) | −0.480$^{*}$ | (0.257) | [−2.901 | 0.142] |
| *Education* | | | | | | |
| – Years of completed schooling | 0.713$^{***}$ | (0.034) | 0.720$^{***}$ | (0.034) | [0.448 | 0.864] |
| – Health-related study | 0.370$^{***}$ | (0.008) | 0.370$^{***}$ | (0.008) | [0.326 | 0.383] |
| *Controls* | | | | | | |
| – Lottery cat and year | Interacted | | Separable | | Separable | |
| – Age and gender | Yes | | No | | No | |

*Note*: The first column presents results from the specification used in the main text. The second column reports results from a restricted specification where omits age and gender are omitted and lottery category and year of first lottery are added additively. This shows that results in columns (1) and (2) are almost identical. The last column reports the 95% confidence interval implied by the lower and upper bound based on the trimming method proposed by Lee (2009). Applying Lee's method to the specification of the first column results in some very small cells not satisfying the monotonicity condition. Standard errors are heteroskedasticity robust.

$^{*}$ Significance at the 10%-level.
$^{**}$ Significance at the 5%-level.
$^{***}$ Significance at the 1%-level.

**Table A4**
Sensitivity of 2SLS results to sample selection on observables.

| | Without controls for age and gender (1) | | With controls for age and gender (2) | | IPW (3) | |
|---|---|---|---|---|---|---|
| *Health status* | | | | | | |
| – Self-reported | 0.036 | (0.037) | 0.033 | (0.037) | 0.033 | (0.037) |
| *Health behavior* | | | | | | |
| – Overweight | −0.006 | (0.024) | −0.010 | (0.024) | −0.010 | (0.025) |
| – Obese | 0.004 | (0.009) | 0.005 | (0.009) | 0.002 | (0.010) |
| – Underweight | −0.011* | (0.007) | −0.011* | (0.007) | −0.010 | (0.006) |
| – Drinks | −0.518 | (0.330) | −0.680** | (0.319) | −0.572* | (0.344) |
| – Drinks > 14 | −0.021 | (0.013) | −0.026** | (0.012) | −0.020 | (0.014) |
| – Smoker | −0.022 | (0.016) | −0.025 | (0.016) | −0.021 | (0.017) |
| – # Cigarettes | −0.015 | (1.321) | −0.104 | (1.295) | −0.113 | (1.332) |
| – Sport | −0.100 | (0.065) | −0.104 | (0.065) | −0.098 | (0.065) |
| *Health altruism* | | | | | | |
| – Donor | 0.109*** | (0.028) | 0.112*** | (0.028) | 0.115*** | (0.028) |
| *Family situation* | | | | | | |
| – Single | −0.008 | (0.012) | −0.008 | (0.012) | −0.010 | (0.012) |
| – Nr of children | 0.035 | (0.059) | 0.043 | (0.058) | 0.035 | (0.059) |
| *Labor market situation* | | | | | | |
| – ln(hours worked) | 0.088*** | (0.021) | 0.080*** | (0.019) | 0.078*** | (0.020) |
| – 60+hour working week | 0.020 | (0.018) | 0.018 | (0.018) | 0.015 | (0.019) |
| – ln(hourly wage) | 0.071** | (0.033) | 0.074** | (0.033) | 0.069** | (0.034) |
| – Nr of weeks unemployed | −1.002* | (0.538) | −0.968* | (0.532) | −1.064* | (0.559) |
| *Education* | | | | | | |
| – Years of completed schooling | 1.462*** | (0.067) | 1.463*** | (0.067) | 1.475*** | (0.070) |
| – Health-related study | 0.757*** | (0.011) | 0.759*** | (0.010) | -0.762*** | (0.011) |

*Note*: All models (2SLS and Probit) include controls for lottery weight, year of first lottery and interactions of lottery weight and year of first lottery. Column (2) add controls for age and gender. The results in column (3) are from inverse probability weighted (IPW) 2SLS regressions, where the response probability is estimated using a Probit model with full interactions between lottery result and age, gender and lottery category. Standard errors are heteroskedasticity robust.

* Significance at the 10%-level.
** Significance at the 5%-level.
*** Significance at the 1%-level.

effects of the content and length of their education and their higher earnings.

## Acknowledgement

We gratefully acknowledge valuable comments from two anonymous referees, Erik Plug and seminar participants in Amsterdam and Madrid.

## Appendix A.

Tables A1–A4

## References

Albouy, V., Lequien, L., 2009. Does compulsory education lower mortality? Journal of Health Economics 28, 155–168.
Arendt, J., 2005. Does education cause better health? a panel data analysis using school reform for identification. Economics of Education Review 24, 149–160.
Baldwin, P., Dodd, M., Wrate, R., 1997. Young doctors' health - II. Health and health behavior. Social Science and Medicine 45, 41–44.
Braakmann, N., 2011. The causal relationship between education, health and health related behaviour: evidence from a natural experiment in England. Journal of Health Economics 30, 753–763.
Cutler, D.M., Lleras-Muney, A., 2008. Education and health: evaluating theories and evidence. In: Schoeni, R.F., House, J.S., Kaplan, G., Pollack, H. (Eds.), Published in Making Americans Healthier: Social and Economics Policy as Health Policy. Russell Sage Foundation, New York.
Frijters, P., Haisken-DeNew, J., Shields, M., 2005. The causal effect of income on health: Evidence from German reunification. Journal of Health Economics 25, 997–1017.
Grossman, M., 2006. Education and non-market outcomes. In: Hanushek, E., Welch, F. (Eds.), Handbook of Economics of Education. Elsevier Science, Amsterdam.
Idler, E., Benyamini, Y., 1997. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior, 21–37.
Imbens, G.W., Angrist, J.D., 1994. Identification and estimation of local average treatment effects? Econometrica 62 (2), 467–475.
Kemptner, D., Jürges, H., Reinhold, S., 2011. Changes in compulsory schooling and the causal effect of education on health: evidence from Germany. Journal of Health Economics 30, 340–354.
Ketel, N., Leuven, E., Oosterbeek, H., Van der Klaauw, B., 2012. The return to medical school in a regulated labor market: evidence from admission lotteries, unpublished working paper.
Kirkcaldy, B., Furnham, A., Shephard, R., 2008. The impact of working hours and working patterns on physical and psychological health. In: Cartwright, S., Cooper, C.L. (Eds.), The Oxford Handbook of Organizational Well Being. Oxford University Press.
Lee, D.S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. Review of Economic Studies 76, 1071–1102.
Lindahl, M., 2005. Estimating the effect of income on health and mortality using lottery prizes as exogenous source of variation in income. Journal of Human Resources 40, 144–168.
Lleras-Muney, A., 2005. The relationship between education and adult mortality in the united states. Review of Economic Studies 72, 189–221.
Oreopoulos, P., 2007. Do dropouts drop out too soon? Wealth, health, and happiness from compulsory schooling. Journal of Public Economics 91, 2213–2229.
Park, C., Kang, C., 2008. Does education induce healthy lifestyle? Journal of Health Economics 27, 1516–1531.
Ross, C.E., Mirowsky, J., Goldsteen, K., 1990. The impact of the family on health: The decade in review. Journal of Marriage and the Family 52, 1059–1078.
Schmitz, H., 2011. Why are the unemployed in worse health? The causal effect of unemployment on health. Labour Economics 18, 71–78.
Tenn, S., Herman, D.A., Wendling, B., 2010. The role of education in the production of health: an empirical analysis of smoking behavior. Journal of Health Economics 29, 404–417.
Webbink, D., Martin, N.G., Visscher, P.M., 2010. Does education reduce the probability of being overweight? Journal of Health Economics 29, 29–38.
Wilson, C., Oswald, A., 2005. How does marriage affect physical and psychological health? A survey of the longitudinal evidence. IZA DP No. 1619.
Wood, R.G., Goesling, B., Avellar, S., 2007. The effects of marriage on health: A synthesis of recent research evidence. Technical report, Mathematica Policy Research, Inc.