The returns to medical school in a regulated labor market: Evidence from admission lotteries[*]

Nadine Ketel[†] Edwin Leuven[‡]

Hessel Oosterbeek[§] Bas van der Klaauw[¶]

Abstract

We estimate the returns to medical school by exploiting that admittance to medical school in the Netherlands is determined by a lottery. Using data from up to 22 years after the lottery, we find that in every single year after graduation doctors earn at least 20 percent more than people who end up in their next-best occupation. Estimated earnings profiles suggest that the life-time difference is even much larger: 22 years after the lottery the earnings difference is almost 50 percent. Only a small fraction of this difference can be attributed to differences in working hours and human capital investments. We therefore interpret the return as a rent due to the restricted supply of doctors in the Netherlands. The returns do not vary with gender or ability, but are lower for individuals with a stronger preference for medical school.

JEL-codes: J44; I18; I23; C36

Keywords: Medical school, returns to education, occupational licensing, random assignment

[*]This version: March 2013. We gratefully acknowledge valuable comments from Maarten Lindeboom, Sarah Turner, Dinand Webbink, and seminar participants in Amsterdam, Bern, Bonn, Bristol, Lausanne, Munich, Oslo, The Hague and Toulouse.

[†]VU University Amsterdam; University of Amsterdam; Tinbergen Institute

[‡]University of Oslo. Also affiliated with CEPR, CESifo, IZA and Statistics Norway.

[§]University of Amsterdam; TIER

[¶]VU University Amsterdam; Tinbergen Institute; CEPR; IZA

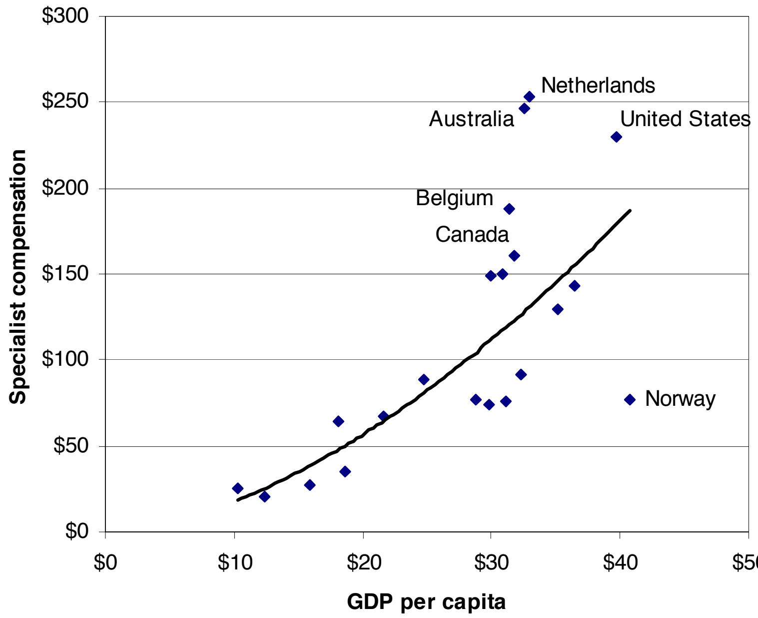
1 Introduction

This paper exploits admission lotteries for medical school in the Netherlands to estimate the earnings return to completion of medical school. Like in many other countries, the supply of doctors in the Netherlands is restricted by limiting the number of places in medical school (Simoens and Hurst, 2006). Our approach produces credible estimates of the rents that doctors earn in such a regulated labor market.

Restricting the supply of doctors can be motivated on the grounds of (i) ensuring the quality of the study program, (ii) the high costs of medical studies which would be wasted for students who don't find employment as a doctor, and (iii) avoiding supply-induced demand. Knowledge about the size of the rents that doctors earn, allows governments to take informed decisions about the tightness of the supply restriction, the level of tuition fees, and regulations regarding doctors' earnings levels.

Previous studies that looked at the monopoly rent of doctors have either ignored selectivity issues or examined very specific groups. Friedman and Kuznets (1954) quantify the rent for doctors in the 1950s in the US by comparing earnings of doctors to earnings of dentists, for whom at the time entry was much less restrictive. They find that 16.5 percent of doctors' earnings is due to "barriers to entry". They admit, however, that part of the observed earnings difference may reflect ability differences. Burstein and Cromwell (1985) follow the same approach and find that in 1978-1980 in the US, the income difference between doctors and dentists amounts to 35 percent, while the income difference between doctors and lawyers is 139 percent. More recently, also using US data, Anderson et al. (2000) show that doctors in states with higher entry barriers due to stricter regulations earn significantly higher incomes. Finally, Kugler and Sauer (2005) measure the effects of licensing by exploiting a retraining assignment rule for immigrant doctors in Israel. They find that immigrant doctors have monthly earnings that are on average 180 percent to 340 percent higher after obtaining a license. These returns only apply to immigrants who have at least 20 years of experience as a doctor.

We use administrative data from the admission lotteries in the years 1988 to 1999, and of applicants' subsequent study career from the Dutch student registry. This information



Source: Congressional Research Service (CRS) analysis of Remuneration of Health Professions, OECD Health Data 2006 (October 2006), available at [http://www.ecosante.fr/OCDEENG/70.html].

Figure 1. Specialist Compensation and GDP per Capita (in U.S. $1,000s), 2004

is merged at the individual level with data on labor market outcomes in the period 1999 to 2010. For the cohort that applied to medical school in 1988, we thus have labor market information of up to 22 years after their application. We present separate estimates for each year since the first application thereby constructing synthetic experience-earnings profiles.

The admission lotteries that we exploit in this paper allow us to deal with selectivity issues. Because applicants who lose the lottery are allowed to reapply in the next year and because not all lottery winners complete medical school, we use the outcome of an individual's first lottery as instrumental variable for completing medical school. The first stage effect is substantial and highly significant.

Even in the absence of the supply restriction, earnings as a doctor may differ from earnings in the next-best occupation. Two reasons are often mentioned. First, the investments in human capital differ. Becoming a doctor requires, on average, more years of education than necessary to enter other occupations. Because we observe the entire earnings profile since first applying to medical school and tuition fees are known, we can address these differences. Second, job characteristics differ between professions, most importantly, it is often claimed that doctors work long hours. We examine this possible explanation for the earnings differential by separately analyzing differences in working hours.

We find substantial earnings returns to completing medical school. There is no single year after graduation in which the returns are less than 20 percent. The earnings profiles indicate that returns increase with experience. Twenty-two years after the first lottery doctors have, on average, almost 50 percent higher earnings. The returns are very similar for men and women, although in absolute terms men earn more than women. We also do not find differential returns by ability, where ability is measured by high-school GPA. Using the approach developed by Imbens and Rubin (1997), we analyze the marginal distribution of earnings under different treatments. This reveals that the large earnings returns are not only driven by high returns in the top end of the distribution. There are fewer people with zero earnings and the whole earnings distribution is shifted to the right.

The large earnings differences cannot be attributed to differences in working hours. While doctors work longer hours than non-doctors, this difference is modest. In the first four years of their careers doctors work annually around 300 hours more. After this first period this difference shrinks to around 100 hours per year. There is also no evidence that doctors are more restricted in their private lives. Doctors are more likely to be married and to have children.

The instrumental variable approach implies that we identify average treatment effects for applicants who complete medical school if they win their first lottery, and who do not complete medical school if they lose the first lottery. Not entering medical school after losing the first lottery can result from not reapplying or from losing subsequent lotteries. We exploit information about participation in and outcomes of subsequent lotteries to further characterize the compliers. In particular, we show that we can identify separately the earnings returns for compliers by the maximum number of lotteries in which they are willing to participate. We interpret the latter as an indicator of the preference to work in the medical profession. We find that the earnings returns are largest for compliers who only want to participate in a single admission lottery. The earnings returns decrease in the maximum number of times a complier is prepared to participate in an admission lottery. This is due to those with the lowest preferences having worse outcomes in their next-best occupations, indicating that preferences may also be correlated to more general motivation for working.

A possible confounding factor in the instrumental variable strategy is that disappointment of losing the first lottery may have a direct effect on earnings in the next-best occupation. To assess this possible channel, we report results from admission lotteries for two other university studies, one of which does not give access to a regulated labor market. We find no earnings difference between winners and losers of the first lottery for that study. This suggests that the disappointment of losing an admission lottery does not reduce future earnings.

The remainder of this paper is organized as follows. The next section provides further details about the institutional context and the admission lottery to medical school. Section 3 describes the data used in this paper. Section 4 discusses the empirical model and the identification. Section 5 presents the estimation results. Section 6 discusses the interpretation of our results in terms of characterizing different types of compliers. Section 7 concludes.

2 Background and institutional context

2.1 Medical schools in the Netherlands

In the Netherlands students choose their field as soon as they enter university, unlike, for example, in the US where students specialize later. Graduates from the pre-university track in high school can enroll in all fields at all universities.[1] Universities have to accept all applicants. However, some fields have a quota, implying that only a fixed number of students is admitted.

The quota for medical schools was introduced in 1976. Initially, the argument for the quota was to ensure the quality of the study program in a time of increasing numbers of

[1]Students are tracked into different levels when they enter high school at age 12. Only the highest of three levels ensures direct admittance to university. Around 20 percent of primary school students enroll in the highest track.

applicants. More recently, the arguments in favor of the quota are threefold (RVZ, 2010). First, since university education is largely publicly funded and medical school is much more expensive than the average study, it is considered a waste of resources to educate doctors for whom there is no employment as a doctor. Second, the teaching capacity of medical schools is limited, and increasing the capacity might reduce the quality. Finally, there may be supplier induced demand (Hurley, 2000), implying that educating more doctors increases the number of medical treatments.

The minister of education decides about the size of the quota. Until 1993 the annual quota was fixed at 1458 students. From 1993 to 1995 it was gradually expanded to 1815 students in 1995. In the years relevant for our paper it remained at this level. The size of the quota is based on the number of places in specialization tracks, which is determined by the associations of specialists. For example, the association of neurologists decides how many places there are available for the specialization tracks in neurology.

If the number of applicants for medical schools exceeds the quota (which has always been the case), a lottery determines who is admitted.[2] Rejected applicants are allowed to reapply in the next year, and until 1999 they could do this as often as they wanted.[3] We observe that 69 percent of the rejected first-time applicants reapply a second time.

The lottery is weighted such that students with a higher GPA on the secondary school exams have a higher probability of being admitted.[4] High school exams are nationwide and externally graded on a scale from one to ten, where six and above indicates a pass. Table 1 shows which GPA intervals are assigned to the different lottery categories - labelled A to F -, together with the shares of applicants in each category. The final column indicates the weights in the lottery. This weight determines the ratio of places assigned to a category over the number of applicants in this category relative to category D. Hence, someone in category A has a twice as high probability of being admitted than someone in category

[2]Since 2000, medical schools are allowed to admit at most 50 percent of the students using their own criteria. Medical schools have made increasing use of this, and selection is often based on motivation and previous experience.

[3]In our data, the maximum number of applications of one individual is nine. In 1999, the maximum number of applications was limited to three.

[4]Graduating from high school requires an exam in seven subjects including Dutch and English. Applicants for medical school should also have included biology, chemistry, physics and math and should have passed these subjects.

Table 1. Lottery categories

Category	GPA	Share	Weight
A	$GPA \geq 8.5$	0.02	2.00
B	$8.0 \leq GPA < 8.5$	0.05	1.50
C	$7.5 \leq GPA < 8.0$	0.09	1.25
D	$7.0 \leq GPA < 7.5$	0.21	1.00
E	$6.5 \leq GPA < 7.0$	0.22	0.80
F	$GPA < 6.5$	0.30	0.67
Other	-	0.11	1.00

Notes: GPA is grade point average on the final exams in high school. Share is the share of applicants in the different categories that applied for the lotteries in the years 1988-1999. Weight indicates the relative probability of being admitted. The category other refers to students who did not participate in the nationwide high school exams, such as foreign students. This category will be excluded from the analysis

D.[5]

Figure 2 shows the admission rates by lottery category.[6] In the early years applicants in category A are almost certainly admitted but this category contains only 2 percent of all applicants (see Table 1). The majority of applicants are in categories C to F, for which the admittance rates range from 35 to 60 percent. Since applicants can participate in multiple lotteries, eventually almost 72 percent of all first-time applicants between 1988 and 1999 are admitted.[7]

The admission lottery is centrally executed. Applicants are allowed to list their first three preferred medical schools. After the result from the lottery is known, admitted students are divided over the medical schools taking account of their preferences where possible. In the Netherlands, eight universities have a medical school, which offer programs that are similar in content and quality. Universities are publicly funded and the nationwide tuition fee is low and fixed by the government. There are no private institutes offering

[5]The total number of available places are divided over categories A to F such that for the number of available places divided by the number of applicants in a category, the weights as in Table 1 hold. In case the number of available places in a category exceeds the number of applicants, all applicants in that category are admitted. For the remaining categories the weights between the ratios of available places and the number of applicants per category will remain the same.

[6]Table A1 in the appendix contains more detailed information on the admission probabilities together with the number of applicants per category per year.

[7]In 1999 a reform was implemented which implied that from that year on applicants with a GPA above 8 (category A and B) are automatically admitted. This reform was implemented as a response to a large public discussion about a girl that finished high school with an exceptional GPA of 9.6 but lost the lottery three times in a row. The weights for the other categories remained the same.

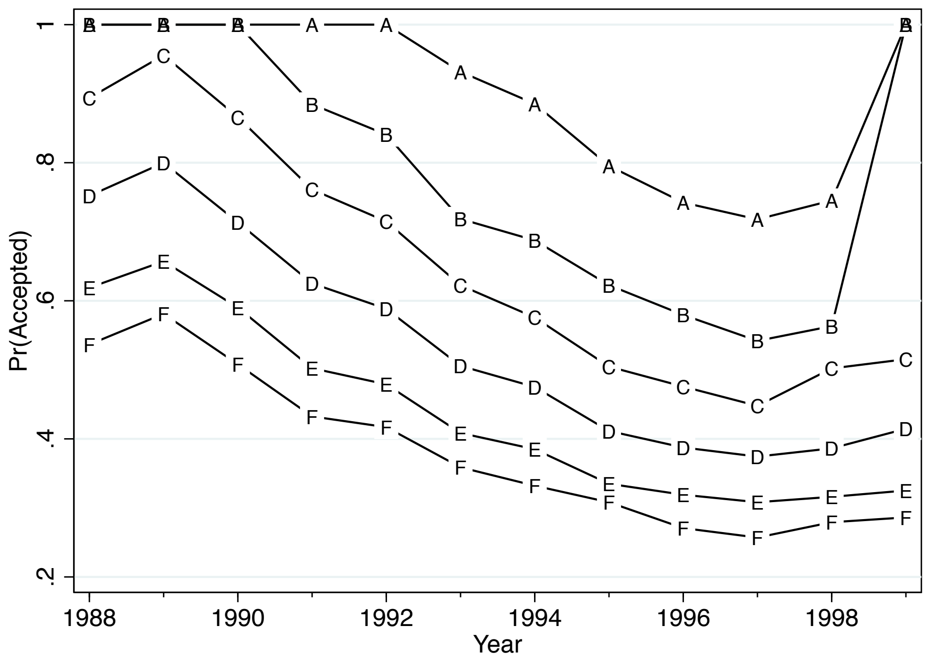


Figure 2. Probability of being admitted by year of application

the same education.

The study program of medical schools consists of different phases. After completing four years of mainly theoretical education students receive their undergraduate diploma. To enter the labor market for medical doctors, two more years of practical training are required. After that, students can choose to specialize. The specialization for a general practitioner takes three additional years while the most advanced specializations such as neurologist, cardiologist or surgeon require an additional four to six years of training. In order to get a place in one of the medical specializationn tracks it is common to first get a PhD degree. In total, the complete medical study can take between six and 15 years.

During the first six years students are entitled to the same general study allowance that all Dutch students receive and students pay a tuition fee of around 1000 euro (at that time). During the PhD or specialization track students are not charged a tuition fee. Instead, they have a formal employment contract and receive a salary.

2.2 The labor market for doctors in the Netherlands

On average, 45 percent of all licensed doctors in the Netherlands are registered as a specialist.[8] General practitioners constitute around 30 percent of the physician population. The remainder pursues a career as a social doctor (10 percent)[9] or does not specialize at all (15 percent). The latter group can either be non-specialized doctors that work as a so-called "basisarts" or those that completed a medical degree, registered as a health care professional but are no longer active as a doctor. There are gender differences in the career choices of doctors. Men are more likely than women to become a specialist (52 versus 39 percent) and less likely to become general practitioner (25 versus 31 percent), social doctor (8 versus 10 percent), or to never specialize (15 versus 20 percent).

A medical specialist can either become an employee of a hospital or can join a medical partnership, which is a joint venture of self-employed individuals. Within hospitals, most specialists (75 percent) are organized in such partnerships (Schafer et al., 2010). Members of a partnership are considered to be self-employed and are taxed as such. The hospital buys the services of these partnerships. During our observation period (1999-2009) two payment regimes applied for self-employed specialists. From 1999 to 2005, each partnership received a lump-sum payment, negotiated by local initiatives of partnerships, insurance companies and hospitals. After 2005, the lump-sum payments were combined with fees per service, in order to introduce incentives for providing services.

The number of practicing doctors in the Netherlands is 2.9 per 1000 inhabitants, close to the OECD average of 3.1.[10] Also the division amongst general practitioners, specialists and other doctors in the Netherlands is close to the OECD average. The same holds for the number of medical graduates; in 2009, 9.9 per 100,000 inhabitants. In terms of remuneration general practitioners in the Netherlands earn 1.7 or 3.5 times the average income depending on whether they are regular employees or self-employed. Self-employed GP's in the UK, Ireland, Germany and Canada have comparable relative remuneration

[8]In order to practice as a physician a doctor needs to be registered in the Dutch registration of health care professionals.

[9]The category of social doctors includes, for example, occupational health doctors, doctors for mentally disabled, community doctors, etc.

[10]The information in this paragraph comes from OECD (2010).

rates. Specialists in the Netherlands are well paid at 5.5 times the average income. There is no other country where this ratio is so high, although also in several other countries (including Australia, Austria, Canada, Ireland and Germany) this ratio exceeds four.

Doctors with a non-Dutch diploma can practice in the Netherlands if their diploma is recognized by the Dutch registration authority. They often have to follow a number of years of additional training, depending on the assessment of their diploma. In the period 2000-2004, 191 non-EU doctors obtained a medical degree following this procedure (Herfs, 2009). Since 2005 non-EU citizens also have to pass a language test and a medical ability test. The language tests are a considerable barrier; in the years 2005-2009 only 19 participants (one quarter of all participants) passed the tests (Herfs, 2009). For EU-citizens the Dutch government is not allowed to demand a language requirement, but employers can. In practice, many employers ask candidates to pass the same language test as non-EU citizens. There are no exact numbers on the number of foreign doctors practicing in the Netherlands. By linking information from the Dutch registration authority to study registrations we observe that 94 percent of the licensed doctors born after 1970 attended medical school in the Netherlands. The remaining six percent will be a combination of Dutch students that attended medical school abroad and foreign doctors that registered in the Netherlands.

3 Data

3.1 Data sources and sample

Our data come from two sources. The first source are the administrative records from the agency (DUO) that registers enrollment of all Dutch students in higher education and that conducts the admission lotteries. Hence, we observe all applicants for medical school together with their lottery category and the outcomes of the lotteries. Furthermore, we know the actual study choices of both winning and losing lottery applicants. Information on study progress is also available as the agency registers when and whether students successfully complete certain stages.

We have lottery data of individuals that applied for a lottery between 1987 and 2004.

Because we are interested in the full history of lottery participation, we exclude individuals who participated in the lottery in 1987. For that year, we cannot observe if participation in the lottery is preceded by losing previous year's lottery. Our data show that people very rarely skip lottery years. But if someone applied in 1986 and next in 1988 (so skipped 1987), we would mistakenly consider the lottery participation in 1988 as start of the application history. To minimize such mistakes, we exclude applicants that are older than 20 at the moment of their first observed application.[11] Since 2000, Dutch medical schools can admit at most 50 percent of their students using their own criteria. Therefore, we exclude all applicants that applied for the first time after 1999. Finally, applicants in lottery category A are excluded since almost all of them are eventually admitted to medical school. This leaves us with 25,551 individuals.

Using social security numbers, the information from DUO is merged to individual records of all Dutch citizens kept by Statistics Netherlands. We lose 60 observations without a valid social security number, which are evenly distributed among the winners and losers of the first lottery (p-value of equality is 0.18). The records of Statistics Netherlands include information from municipalities, tax authorities and social insurance administrations. Therefore, they contain detailed information on earnings from various sources, labor supply and characteristics such as age, gender, ethnicity and marital status. All inhabitants of the Netherlands are registered at a municipality, which implies that if a person is not in our data in a particular year, this person did not live in the Netherlands in that year. Data from Statistics Netherlands cover the years 1999 to 2010, with the exception of working hours which are only available for the years 2006-2010. Finally, we have records from the BIG-register, that includes all health care professionals in the Netherlands. This register provides information regarding individual qualifications and entitlement to practice. From this register we know whether someone is licensed as a doctor.

[11]In the Netherlands, the nominal age of finishing high school is 18.

Table 2. Descriptive statistics by admission status of the first lottery application

	Lottery winners	Lottery losers
Personal characteristics		
Female	58%	58%
Age at first application	18.3	18.3
Non-western immigrant	8%	8%
GPA high school exams (0-10)	7.06	6.79
Study enrollment and completion		
Enrolled in medical school	94%	45%
Completed medical school (by 2010)	83%	41%
Licensed as doctor (by 2008)	80%	42%
Enrolled in study program in the Netherlands	99%	95%
Completed study program in the Netherlands	96%	89%
Labor market outcomes		
Annual inflation-adjusted (to 2010) taxable earnings (1999-2010)	39,149	28,268
Annual working hours (2006-2010)	1756	1693
Hourly earnings (2006-2010)	31.5	24.1
Ever collected any social insurance benefits	18%	19%
Household composition		
Married in 2010	51%	45%
Children in 2010	60%	51%
Number of individuals	11,819	13,672

Note: Recall that the lottery is weighted so that the observed differences between lottery losers and lottery winners cannot be given a causal interpretation.

3.2 Descriptive statistics

Table 2 presents descriptive statistics separately for winners and losers of their first lottery.[12] The upper part of the table provides information on personal characteristics. The majority of the applicants is female and the percentage of women is similar among winners and losers. The average age at the first application is 18.3. The mean GPA of lottery winners is higher than of lottery losers, which reflects that GPA determines the weight in the lottery.

Next, the table presents summary statistics on study enrollment and completion. The outcome of the first lottery is associated with almost a 50 percentage point increase in enrollment into medical school. But not everyone who wins the first lottery actually enrolls in medical school; six percent do not enroll. Among the losers of the first lottery

[12] When there can be no confusion we sometimes refer to winners and losers of their first lottery as "lottery winners" and "lottery losers".

45 percent enroll in medical school (after winning a subsequent lottery). Of the winners 83 percent complete medical school, and this is 41 percent for the losers. Finally, almost all individuals who complete medical school also register as a doctor, and are therefore licensed. The small difference between completion and registration rates may be caused by the fact that completion data currently run to 2010 while the register of licensed doctors only runs to 2008. Additionally, for lottery losers it might be that some individuals obtained a medical degree abroad and afterwards registered as a doctor in the Netherlands.[13]

For the interpretation of the estimated returns to medical school it is important to know which alternatives the lottery losers choose. The majority of the lottery losers attends a study program in the Netherlands.[14] Only five percent of the lottery losers never register for higher education in the Netherlands. These individuals may not have enrolled in any study program or may have studied abroad. Of the lottery participants that do not enroll in medical school but do enroll for Dutch higher education 32 percent enroll in a health related field. Other regularly chosen fields are Sciences (15 percent), Social and Behavioral Sciences (15 percent), Engineering (ten percent), Economics (nine percent) and Law (six percent).

Lottery losers are seven percentage points less likely to complete a study program. This may be due to the fact that medical schools have much lower dropout rates than other study programs. It is often argued that this is the consequence of the intensity of the study program at medical school (more workgroup classes and fewer exams). Also the fact that lottery losers have, on average, lower ability (GPA), may explain their lower graduation rate. This explanation is supported by results from a regression of having a diploma on GPA: Applicants in lottery category F are seven percentage points less likely to obtain a diploma than applicants in lottery category B.

The lower part of Table 2 shows descriptive statistics on several labor market outcomes. We focus on the following outcomes: earnings, working hours, hourly earnings

[13]The Netherlands shares its southern border and its language with the northern part of Belgium, which does not have a quota for medical schools. Instead, applicants have to pass an entry exam.

[14]Recall that enrollment for almost all study programs in the Netherlands is unlimited and unrestricted.

and collecting (social insurance) benefits. Earnings are measured as the sum of before-tax income from employment, income from self-employment, income from abroad and other income from labor. Earnings are observed annually. All amounts are corrected for the average wage development of university graduates over the observation period and converted to 2010 euros. Earnings are set equal to zero for people who live in the Netherlands and have no income from labor.[15] This also includes students without earnings. Table 2 shows that earnings are, on average, around 38 percent higher for winners than for losers.

Information on working hours is only available for 2006 to 2010 and only for employed workers. We assume that self-employed workers have a full-time job (1872 hours per year).[16] Average working hours are close to 1700, but winners work approximately four percent more than losers. This difference is not sufficient to equalize hourly earnings; these are about 31 percent higher for winners. Among both the winners and the losers, about 18 percent ever received some type of social insurance benefits (welfare, unemployment and disability insurance) during the period 1999-2010.

Finally, the bottom part of the table shows descriptives for the household situation in 2010. Winners of the lottery are more likely to be married and to have at least one child.

4 Empirical approach

To estimate the return to medical school we assume a linear relationship between the labor market outcome of individual i in year t who applied for the first time to medical school in year τ ($Y_{it\tau}$) and having completed medical school (D_i):

$$Y_{it\tau} = \alpha_t + \gamma_{t-\tau} + \delta_{t-\tau} D_i + X_i \beta_{t-\tau} + LC_{i\tau} + U_{it\tau} \tag{1}$$

where $t-\tau$ indicates the number of years elapsed between the year of the first lottery and the year in which the outcome is observed. X_i is a vector of controls including gender,

[15]The fraction of people that live abroad increases over time and is five percent in 2008, this fraction is the same for winners and losers.

[16]In case a person has income both from employment and from self-employment we take a weighted average: hours worked = hours from employment + (income from self-employment/total income) * 1872 hours.

ethnicity and age at first lottery, and $LC_{i\tau}$ is the interaction between lottery category and year of first lottery. α_t and $\gamma_{t-\tau}$ are fixed effects for year in which the outcome is observed and the number of years since the first application. $U_{it\tau}$ is the error term. The parameters of interest are $\delta_{t-\tau}$ which describe the returns to medical school $t-\tau$ years after first applying. We estimate equation (1) separately for each number of years since the first lottery participation $(t-\tau)$.

If high-ability students self-select into medical school, the OLS estimator of $\delta_{t-\tau}$ will be biased. The lottery seems to solve this problem, but completing medical school remains potentially endogenous. Not all admitted students actually complete medical school, and lottery losers often reapply in subsequent years. Therefore, we instrument D_i with the result (0/1) of the first lottery (LR_{1i}) in which individual i participated. We estimate a first-stage equation of the form:

$$D_i = \kappa_{t-\tau} + \lambda_{t-\tau} LR_{1i} + X_i\theta_{t-\tau} + LC_{i\tau} + V_{it-\tau} \tag{2}$$

The identifying assumption is that conditional on X_i and $LC_{i\tau}$ the result in the first lottery is mean independent of $U_{it\tau}$: $E(LR_{1i} \cdot U_{it\tau}|X_i, LC_{i\tau}) = 0$. Recall that in each year within each lottery category all individuals have the same probability of being admitted. This conditional random assignment guarantees that the mean conditional independence assumption holds.

In equation (2) the parameter $\lambda_{t-\tau}$ reflects the difference in the completion rates between winners and losers of the first lottery.[17] It will not equal one for three reasons. First, some winners of the first lottery decide not to enroll in medical school. Second, some winners who enroll, do not complete medical school. And third, losers can still obtain a medical degree if they win a subsequent lottery. An interpretation of $\lambda_{t-\tau}$ is that it describes the fraction of compliers in the data, which are applicants for whom graduating from medical school is determined by the result of the first lottery. In Section 6 we elaborate further on the definition of compliers and the interpretation of our estimated

[17]Because we perform separate regressions for the number of years since the first lottery $(t-\tau)$, we also estimate for each value $t-\tau$ a separate λ.

returns to medical school.

By estimating equation (1) separately for each year following the first lottery, we obtain a picture of the evolvement of the earnings differential during the first 22 years after the first lottery. This period captures the longer study duration in medical schools compared to alternative studies, and thereby an estimate of the opportunity costs of the longer investment in human capital.

5 Results

5.1 Main findings

We perform our regressions separately for year after the first lottery ($t-\tau$), which implies that each regression uses different subsamples. Table 3 reports the estimation results in more detail. The second column reports the number of observations in each regression and shows how this varies across rows. The final row ($t-\tau=22$) is only based on 2010-earnings information of people who first applied in 1988. The penultimate row is based on 2010-earnings information of people who first applied in 1989 and on 2009-earnings information of people who first applied in 1988, and so on. Because the admission lotteries in our sample end in the same year in which the earnings data start (1999), also the estimates in the first row are based on just a single cohort.

The first stage regression describes the effect of winning the first lottery on the probability to complete medical school. The third column reports first-stage estimates. The first-stage estimates are highly significant (the F-statistic is never below 290) and are all close to 0.39. So winning the first lottery increases the probability to complete medical school with around 39 percentage points.

The fourth column of Table 3 presents the instrumental variable estimates of the effect of completing medical school on annual earnings (in thousands of euros). The estimates are also plotted in panel (a) of Figure 3. During the first six years after the first lottery the effect is negative or close to zero. In the first four years the reduction is modest for two reasons. First, students who are not in medical school more often have a small job while studying, and second, some people that are not admitted to medical school will

Table 3. Instrumental variable estimates of the effects of completing medical school on labor market outcomes $t - \tau$ years after first applying

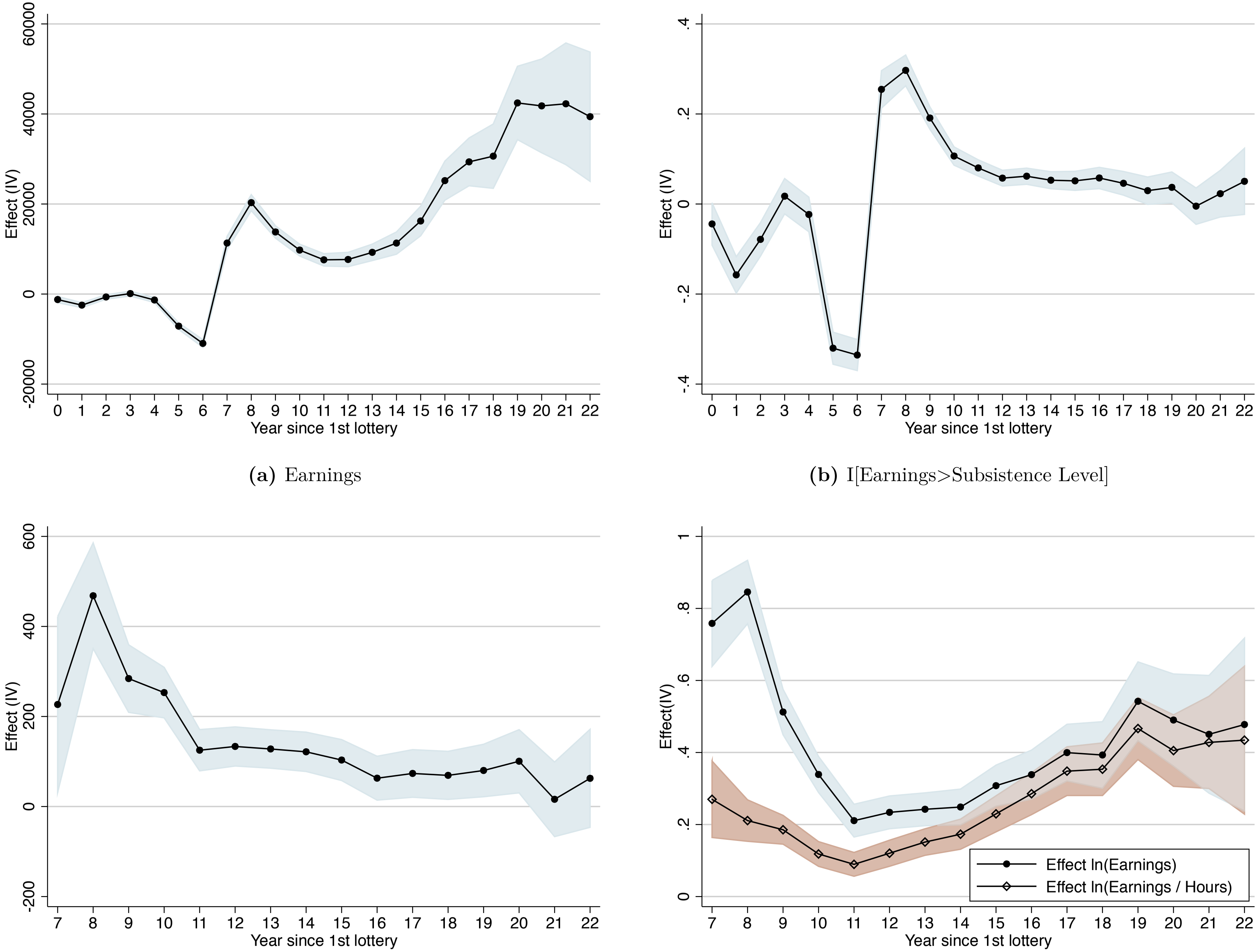
$t - \tau$ (1)	N (2)	1st stage (3)	Earnings (x1000) (4)	I[Earnings>Welfare] (5)	log(Earnings) (6)	Hours (7)	log(Earnings/Hrs) (8)
0	2159	0.36 (0.02)***	-1.2 (0.3)***	-0.04 (0.02)*	-0.36 (0.15)**		
1	4607	0.36 (0.01)***	-2.5 (0.3)***	-0.16 (0.02)***	-0.67 (0.12)***		
2	7167	0.37 (0.01)***	-0.7 (0.3)***	-0.08 (0.02)***	-0.02 (0.08)		
3	9885	0.38 (0.01)***	0.1 (0.3)	0.02 (0.02)	0.07 (0.06)		
4	12,438	0.39 (0.01)***	-1.3 (0.3)***	-0.02 (0.02)	-0.18 (0.06)***		
5	14,952	0.39 (0.01)***	-7.1 (0.4)***	-0.32 (0.02)***	-1.35 (0.07)***		
6	17,154	0.39 (0.01)***	-11.0 (0.4)***	-0.34 (0.02)***	-1.31 (0.07)***		
7	18,945	0.38 (0.01)***	11.3 (0.8)***	0.25 (0.02)***	0.76 (0.06)***	227 (100)**	0.27 (0.05)***
8	20,705	0.38 (0.01)***	20.3 (0.9)***	0.30 (0.02)***	0.85 (0.04)***	468 (60)***	0.21 (0.03)***
9	22,183	0.38 (0.01)***	13.8 (0.7)***	0.19 (0.01)***	0.51 (0.03)***	284 (38)***	0.19 (0.02)***
10	23,484	0.38 (0.01)***	9.8 (0.7)***	0.11 (0.01)***	0.34 (0.03)***	253 (29)***	0.12 (0.02)***
11	24,849	0.38 (0.01)***	7.6 (0.7)***	0.08 (0.01)***	0.21 (0.02)***	125 (23)***	0.09 (0.02)***
12	22,608	0.38 (0.01)***	7.7 (0.8)***	0.06 (0.01)***	0.23 (0.02)***	133 (22)***	0.12 (0.02)***
13	20,117	0.39 (0.01)***	9.3 (0.9)***	0.06 (0.01)***	0.24 (0.02)***	128 (22)***	0.15 (0.02)***
14	17,568	0.39 (0.01)***	11.3 (1.3)***	0.05 (0.01)***	0.25 (0.03)***	121 (22)***	0.17 (0.02)***
15	14,894	0.39 (0.01)***	16.2 (1.7)***	0.05 (0.01)***	0.31 (0.03)***	103 (23)***	0.23 (0.03)***
16	12,410	0.38 (0.01)***	25.2 (2.2)***	0.06 (0.01)***	0.34 (0.03)***	63 (25)**	0.29 (0.03)***
17	9949	0.39 (0.01)***	29.4 (2.7)***	0.05 (0.01)***	0.40 (0.04)***	73 (27)***	0.35 (0.03)***
18	7808	0.39 (0.01)***	30.6 (3.6)***	0.03 (0.02)*	0.39 (0.05)***	69 (27)**	0.35 (0.04)***
19	6056	0.40 (0.01)***	42.4 (4.2)***	0.04 (0.02)**	0.54 (0.06)***	80 (30)***	0.47 (0.04)***
20	4306	0.41 (0.02)***	41.8 (5.3)***	0.00 (0.02)	0.49 (0.07)***	100 (36)***	0.41 (0.05)***
21	2784	0.41 (0.02)***	42.3 (6.9)***	0.02 (0.03)	0.45 (0.08)***	16 (42)	0.43 (0.07)***
22	1430	0.45 (0.03)***	39.4 (7.3)***	0.05 (0.04)	0.48 (0.12)***	63 (56)	0.43 (0.10)***

Notes: Standard errors in parentheses. Total number of individuals is 25,491. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Every cell in this table represents a separate regression, which include controls for gender, ethnicity, age in the first lottery year, lottery category, year of first lottery and interaction terms of the year of first lottery and lottery category.

decide to work rather than to study. In the fifth and the sixth year after first applying the reduction in earnings is more substantial. This reflects that most alternative studies have a shorter duration than the six years required for medical school. Individuals who do not attend medical school enter the labor market earlier and start receiving income earlier than individuals attending medical school. These negative earnings effects in the fifth and sixth years express the opportunity cost of the larger investment in human capital of people who complete medical school.

The picture reverses from the seventh years onwards. This is the moment that students from medical school have graduated and start earning, either in the labor market or while being employed in a specialization track. From then on the returns to medical school are always positive and significant. After a big jump in years seven and eight the earnings differential remains positive but decreases until the twelfth year; the 20,000 euro per year difference in the eighth year reduces to less than half of that in the twelfth year. During that period many students from medical school are in specialization tracks. Starting wages in specialization tracks are relatively high but hardly rise while being in the track. From the twelfth year onwards, students from medical school might have finished their specialization track and can start working as a (self-employed) specialist or GP. The earnings difference increases again and eventually amounts to almost 40,000 euro per year in the twenty-second year.

An explanation for the substantial increase in earnings returns in the seventh and eighth year can be found in the fifth column (and in panel (b) of Figure 3). Here the dependent variable is a dummy for earning above the level of welfare benefits. In the fifth and sixth year after the first lottery students in medical school have not yet entered the labor market. Therefore, they are less likely to earn above the level of welfare benefits than those not in medical school. But this reverses in the seventh year after the first lottery. The effect is particularly large seven to nine years after the first lottery. While most students from medical school find (full-time) work immediately after graduating, other students struggle more find stable employment. From year ten onwards medical school graduates are around six percentage points more likely to have earnings above the



(a) Earnings

(b) I[Earnings>Subsistence Level]

(c) Hours

(d) ln(Hourly Earnings)

Figure 3. Instrumental variable estimates of the effects of completing medical school on labor market outcomes $t - \tau$ years after first applying (colored areas are 95 percent confidence intervals).

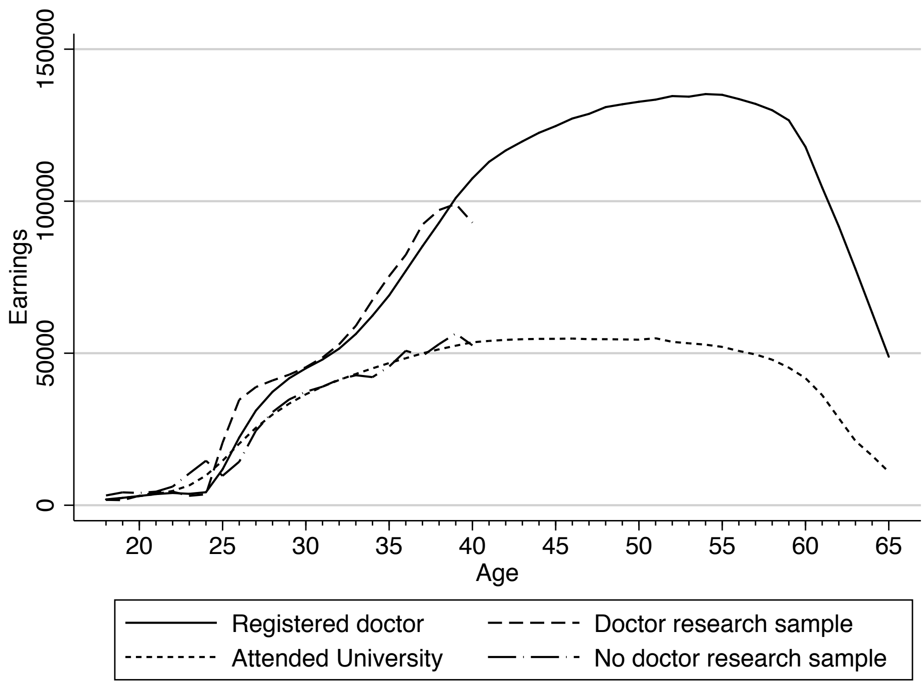


Figure 4. Predicted levels for earnings

level of welfare benefits than other students, relative to a base of 0.89.

In Figure 4 we use the estimated model to show the predicted earnings profiles for an average individual with and without completion of medical school. This implies that for both $D = 0$ and $D = 1$, we compute the expected earnings according to

$$\hat{Y}_{t-\tau} = \hat{\alpha}_t + \hat{\gamma}_{t-\tau} + \hat{\delta}_{t-\tau}D + \bar{X}\hat{\beta_{t-\tau}} + \bar{LC}$$

where $\bar{X}$ and $\bar{LC}$ are the observed sample means. In the figure we assume that individuals first apply to medical school at age 18.

To get an idea how these earnings profiles evolve until retirement, we also plot rescaled average wage profiles of medical school graduates and other university attenders, where the latter serves as reference group for the "non-doctors".[18] We regressed the observed average wage profile medical school graduates on the predicted wage profile for the years in which we observe both, to obtain the scaling factor. We repeated this exercise for other

[18]We have information on all registered doctors in the Netherlands so we can plot their earnings profile until the (retirement) age of 65. The wage profile for other university attenders is the average wage profile of all people for whom it is registered that they attended university and weighted using sampling probabilities. We do not have information on retirement benefits, so we can only take account of earnings while being active on the labor market.

Table 4. Discounted lifetime rents

Discount rate	All	Men	Women
0.02	1,154,629	1,311,651	781,491
0.05	495,315	560,748	359,694
0.10	145,927	162,324	118,117

Notes: The entries represent discounted earnings at the time of participating in the first lottery. The estimated lifetime rents are for a representative individual.

university attenders and non-doctors. The wage profiles show that while being in medical school, students earn less than if they would have attended another study. At the age of 24, after they have finished medical school, they start earning substantially more and remain to do so for the rest of their career. We can only make causal inferences on the effect of completing medical school up to 22 years after participating in the first lottery, but the fitted earnings profiles suggest that the earnings difference is still increasing in the remaining years of the career.

Using the results from Table 3 we calculate the lifetime benefits of completing medical school. For the first 22 year the estimated differences are used so this takes account of the longer study period for doctors and of the two years of unpaid residencies. We assume that in addition to the 22 years since the first lottery that were already estimated, an average career lasts another 24 more years. For the earnings difference in the remaining years we use the difference between the two rescaled wage profiles from Figure 4. Table 4 presents the present values of lifetime rents for discount rates equal to two, five and 10 percent. Earnings are discounted at the time of participating in the first lottery. We see that even for a high discount rate of ten percent the present value of the earnings gain for doctors exceeds 100,000 euro. At a more moderate discount rate of five percent, this present value is already almost 500,000 euro. At a low discount rate of two percent it even exceeds one million euro.

The sixth column of Table 3 and panel (d) of Figure 3 shows results for the effect of medical school on the logarithm of earnings, conditional on having positive earnings. The observed pattern is very similar to the pattern for the level of earnings (which includes zeros) in column (4). During the first six years after the lottery, medical school graduates have lower log earnings than not admitted students and this reverses in the seventh year. Until the tenth year the pattern is a bit erratic, but from the eleventh year the return steadily increases up to 0.54 in the nineteenth year. From then on it remains stable around 0.50; in the last year covered by the data the return is 0.48.

It is often argued that doctors make longer working hours, which should explain (part of) the earnings return. Column (7) of Table 3 and panel (c) of Figure 3 report estimates where annual working hours is the dependent variable. Information about hours is only available for the years 2006 until 2010, and therefore only for the seventh to twenty-second year after the first lottery. The results reveal that doctors work more hours during the first four years after finishing the initial phase of their study. During these four years doctors work a total of 1200 hours more than non-doctors. The average number of working hours during these four years together is around 6,400 hours, so that doctors work around 20 percent more than those that did not complete medical school. After these first four years doctors work about 100 hours more per year than non-doctors. Compared to a baseline of 1,600 hours this is a six percent difference. Differences in working hours can therefore not explain the large earnings gain to medical school. This is confirmed by the results in the final column of Table 3 (and panel (d) of Figure 3) where log earnings per hour is the dependent variable. The effect on log earnings per hour is only marginally smaller than the effect on log earnings. From the eleventh year onwards the gain in the log of per hour earnings increases to 0.43.

Figure 5 shows predicted levels of hours for an average individual with and without medical school. Based on a 36 hour workweek the amount of yearly hours worked is 1,872. We see that from the seventh to the ninth year the non-doctors work around 1,200 hours per year. The large effect that we find on hours in these years is thus mainly driven by the fact that doctors more often have a (full-time) job. After the tenth year the number of hours stabilizes for both groups. In these years the doctors work only marginally more hours than the non-doctors.

Table 3 shows that with the exception of the first few years of their career, doctors do not work much longer hours than others. We now inquire whether the high earnings of

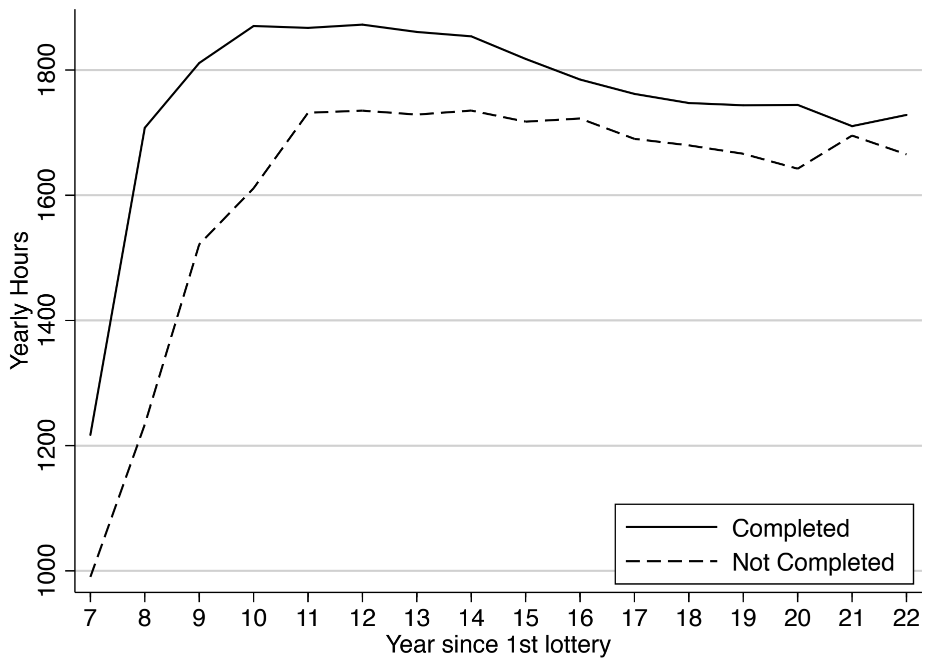


Figure 5. Predicted levels of annual working hours

doctors come at the cost of other outcomes. In Table 5 we look at the impact of medical school completion on having children, being married and ever receiving state benefits (unemployment insurance, disability insurance or welfare). The first column relates to the whole sample, while the second and third columns report results for men and women separately.[19] The results show that doctors never do worse in terms of these outcomes; they do not have fewer children, they are more likely to be married and they are less likely to have ever received state benefits. Results by gender show that male doctors score significantly different on all these variables compared to male non-doctors, while for women the likelihood to be married is not significantly higher for doctors. In short, the household situation of doctors does not suffer from their occupation.

5.2 Heterogenous treatment effects

We now turn to heterogeneity in the returns to medical school. We first examine whether returns differ between men and women. As Table 2 reveals, a majority of the applicants for medical school are female. While of all university students in the Netherlands less than half are female during the period 1988-1999, this is 58 percent in medical schools. This

[19]In the next subsection we report effects on labor market outcomes by gender.

Table 5. Other outcomes

	All	Men	Women
Children in 2010	0.08 (0.02)***	0.12 (0.02)***	0.05 (0.02)***
Married in 2010	0.05 (0.02)***	0.09 (0.02)***	0.02 (0.02)
Ever state benefits	-0.05 (0.01)***	-0.07 (0.02)***	-0.03 (0.02)*

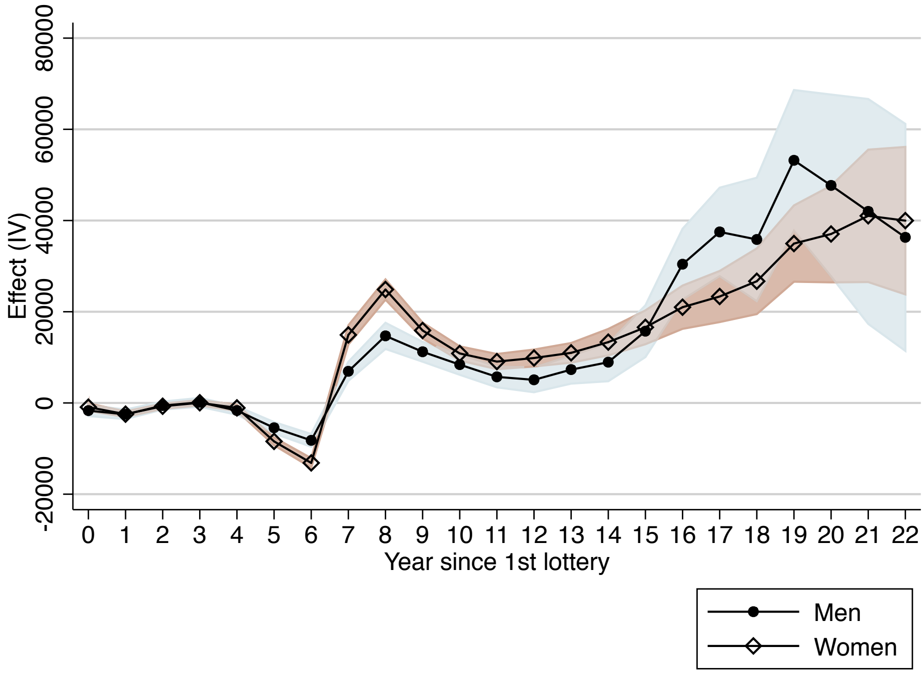
Notes: Robust standard errors in parentheses. Total number of individuals is 25,491. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Every cell in this column represents a separate regression, which include controls for gender (in the first column), ethnicity, age in the first lottery year, lottery category, year of first lottery and interaction terms of the year of first lottery and lottery category

justifies the question whether women have a comparative advantage in medical school. Next we investigate whether returns differ by ability. As described in Section 2 the admission lottery uses weights based on applicants' GPA on secondary school exams. Applicants with a higher GPA have a higher probability of being admitted. This system of a weighted admission lottery justifies the question whether the available places are allocated efficiently. Finally, we study variation in returns over the earnings distribution. This is motivated by the concern often expressed by policy makers, that some medical specializations pay very high wages.

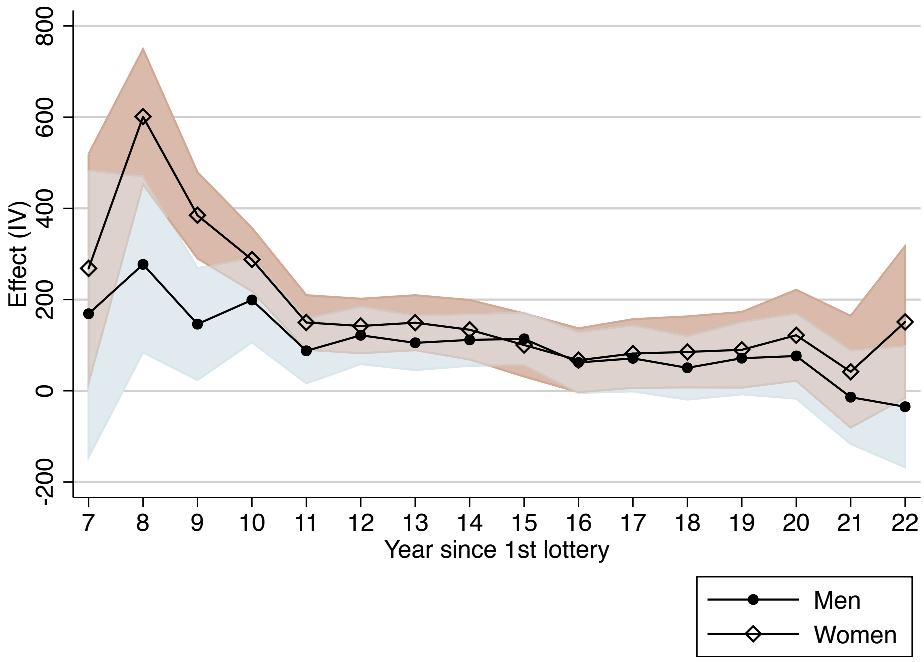
By gender. Figure 6a shows the estimates of the earnings returns separately for men and women.[20] Until the sixth year both men and women experience an earnings loss from studying in medical school. This loss is very similar across gender. In years seven to 14, the returns are much larger for women than for men, but from year 15 onwards men catch up and in the final years the returns to medical school are slightly larger for men than for women. Male doctors more often choose long specialization tracks than female doctors. Figure 7 shows gender-specific predicted earnings profiles with and without medical school. This reveals that female doctors earn more or less the same as male non-doctors, and that doctors earn more than non-doctors of the same sex.

Figure 6b shows that from the seventh to the tenth year after the first lottery, doctors work longer hours than non-doctors of the same gender. The difference is larger for women than for men, although the effects are not significantly different from each other. The effects on hours disappear after the tenth year, and a bit further in their career male

[20]Table A2 in the appendix reports the results.



(a) Earnings



(b) Hours

Figure 6. IV estimates of effects of medical school completion on earnings and hours, by year since first lottery and gender

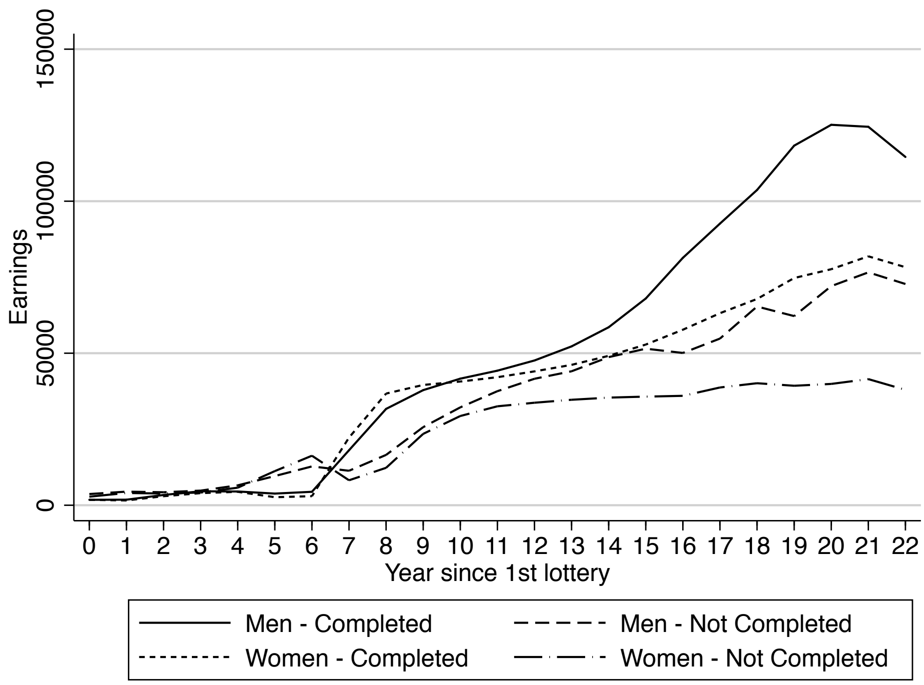


Figure 7. Predicted earnings by gender

doctors even work fewer hours than male non-doctors.

By ability. The lottery gives applicants with higher GPA on their secondary school exams a higher probability to be admitted. This justifies the question whether there is a difference in earnings gain between people with different GPAs. To examine this, we estimated earnings returns by year after first lottery separately for lottery categories B to F.[21] Figure 8 reports the results.[22] The estimates for the early and late years of categories B, C and D are not very precise due to small sample sizes. The results show that the returns are very similar for the different lottery categories with exception of the seventh and eighth year. In these years returns are higher for a higher GPA. This is probably driven by applicants with a higher GPA finishing medical school earlier. If we regress time until diploma on lottery category (conditional on winning the first lottery) we find that students in category F study, on average, half a year longer than students in category B.

As soon as all students have entered the labor market the returns are very similar for the different lottery categories. We do find that the proportion of applicants that

[21]Recall that category A is omitted since there are too few losers in this category.
[22]The results can also be found in Table A3 in the appendix.

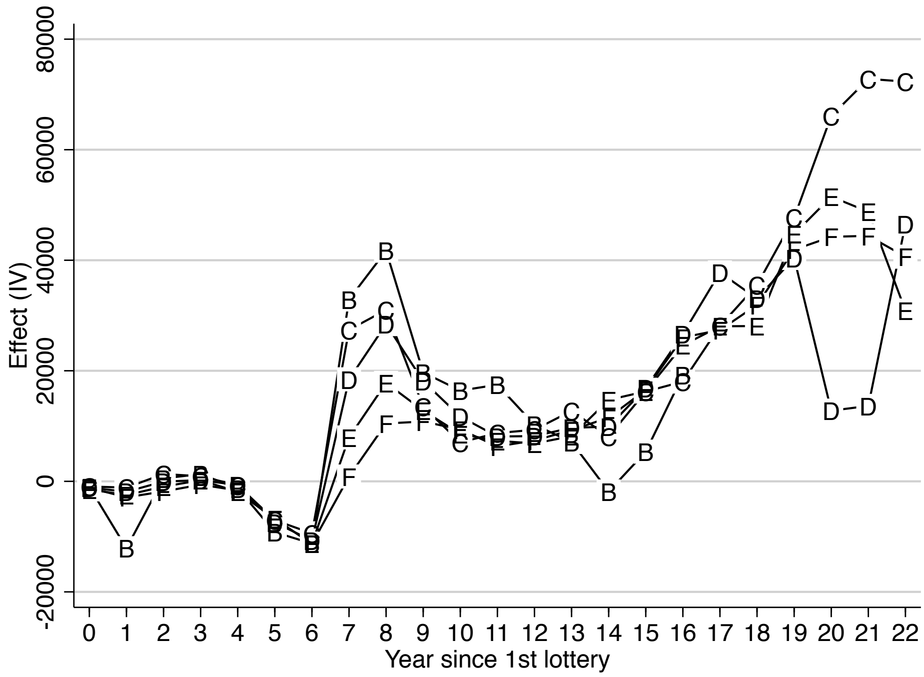


Figure 8. IV estimates of effects of medical school completion on earnings, by year since first lottery and lottery category

becomes a specialist increases with GPA. Conditional on completing medical school, the percentage of specialists decreases monotonically from 59 percent in category B to 42 percent in category F. That we do not find differences in the returns for the different lottery categories follows from the worse outside opportunities that applicants in the lower lottery categories have. Conditional on winning the first lottery, applicants in category F have a 11 percentage points lower probability to complete a degree than applicants in category B. Conditional on losing the first lottery, this difference is 26 percentage points.

If earnings reflect productivity accurately (both in the medical profession and in the second-best professions) and if applicants' GPAs do not respond to changes in the probabilities to be admitted, this implies that there is no clear support for a system in which only students with the highest GPA are selected. The only advantage is that applicants with a high GPA finish their studies faster than students with a low GPA.

Along the distribution of earnings. The common view about the remuneration of doctors in the Netherlands (and elsewhere) is that especially medical specialists are highly paid. The figures about the relative pay of GPs and specialists reported in Section 2, confirm this. This suggests that the earnings gain is distributed unequally across the earnings distribution. To inquire this further we estimate the marginal distribution of the outcome under different treatments for the subpopulation of compliers, following Imbens and Rubin (1997).[23]

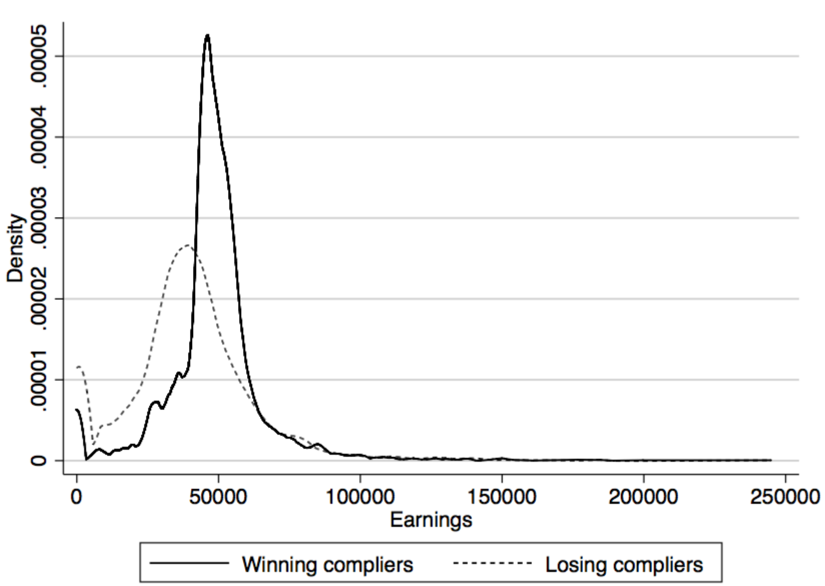
Figure 9 plots the estimated earnings distributions for winning and losing compliers 12, 16, 18 and 20 years after the first lottery. We see that in all these years after the first lottery the distribution of winning compliers has less mass at low incomes than the distribution for losing compliers. After 12 years there is very little dispersion in winning compliers' earnings. This is the time when they do their specialization tracks in which wages are fixed. After 16 years most winning compliers will have finished their specialization track, and the density function of the winning compliers has a similar shape but is shifted to the right compared to the losing compliers. After 18 years the earnings of winning compliers are then more dispersed and the right tail of the winning compliers becomes much fatter, which implies that there are more top earners among the winning compliers. This is even more pronounced after 20 years. These figures show that the earnings gains from medical school that we found in Section 5.1 are not only driven by high gains in the top end of the distribution. Among the winning compliers there are always fewer people who have zero earnings and the distribution of winning compliers is to the right of the distribution of losing compliers.

6 Interpretation

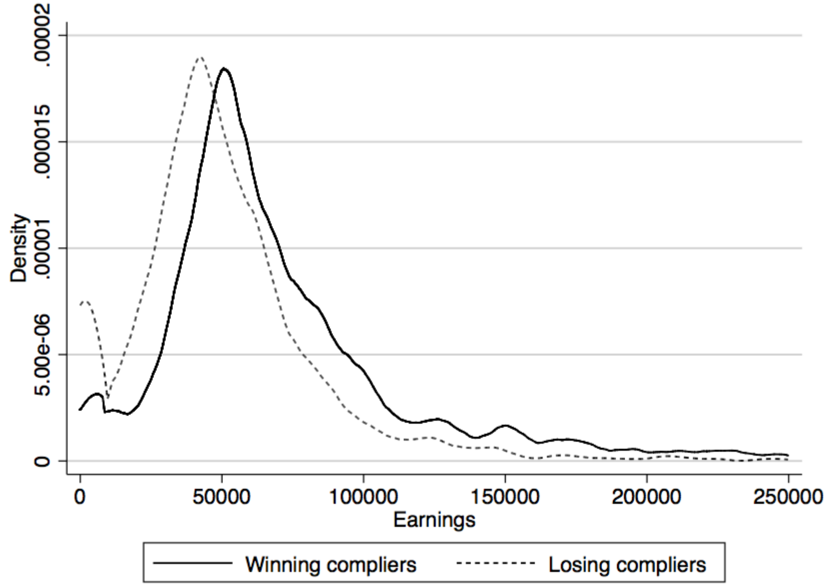
6.1 Degrees of compliance

In this section we elaborate on the interpretation of the instrumental variable estimate in the setting of our analysis. The instrumental variable approach identifies the average returns to medical school for applicants who comply with the result of the first lottery. Following Imbens and Angrist (1994), we can interpret our estimator in terms of potential

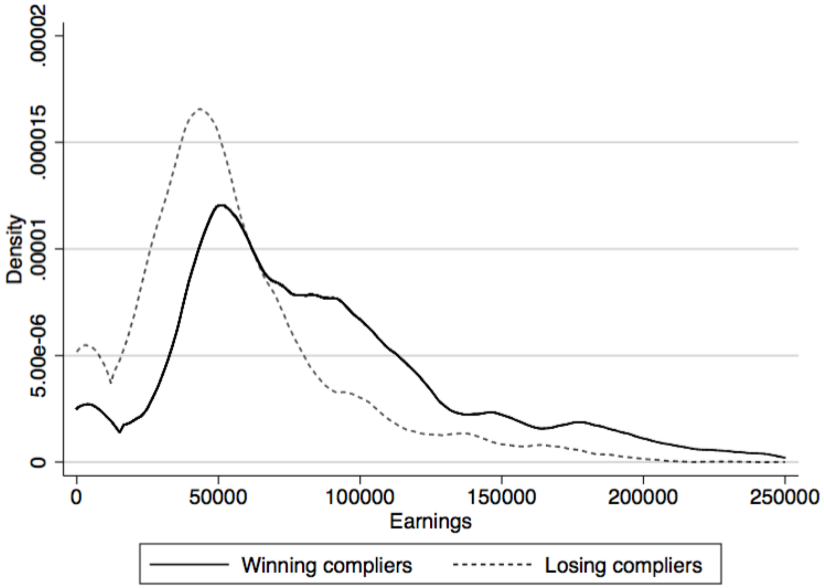
[23]Their method is briefly explained in Appendix B.



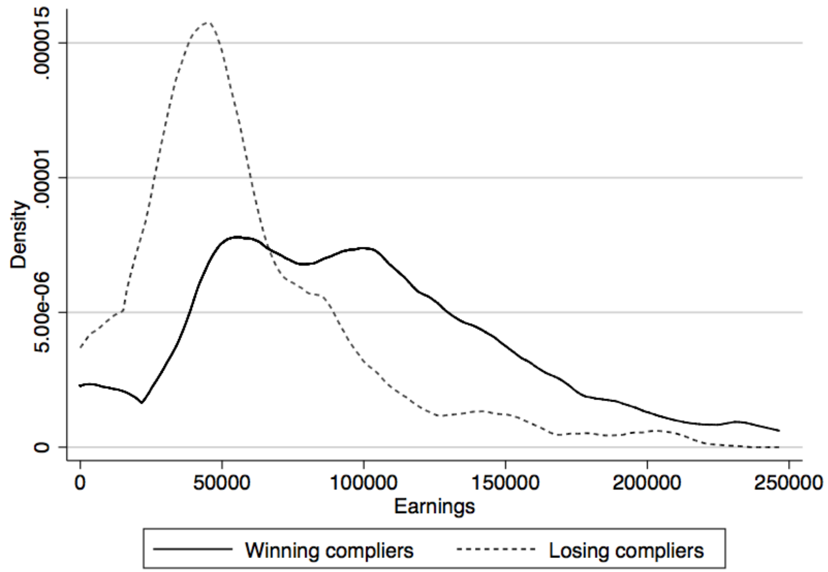
(a) 12 year after the first lottery



(b) 16 year after the first lottery



(c) 18 year after the first lottery



(d) 20 years after the first lottery

Figure 9. Earnings distribution of winning and losing compliers

outcomes:

$$\frac{E[Y|LR_1 = 1] - E[Y|LR_1 = 0]}{P(D = 1|LR_1 = 1) - P(D = 1|LR_1 = 0)} = E[Y(1) - Y(0)|D(LR_1 = 1) - D(LR_1 = 0) = 1] \quad (3)$$

where $Y(1)$ and $Y(0)$ are potential outcomes with and without medical school, respectively, and $D()$ is an indicator for completing medical school if the expression in parentheses holds.

Recall that individuals who lose an admission lottery can reapply in the next year. Even though participating in the admission lottery is free, engaging in a subsequent lottery might be considered costly since individuals postpone entering the labor market. The subsequent lotteries allow us to characterize the compliers in more detail. This is interesting because it can provide information on how earnings returns interact with (other) preferences for completing medical school. One might expect individuals with high earnings returns and individuals with strong preferences for completing medical school to actually participate in subsequent lotteries after losing an admission lottery. This implies that if preferences and earnings returns are independent, the returns to medical school should be higher for individuals who are willing to participate in multiple lotteries.

So far, we considered compliers to their first admission lottery. These are individuals who complete medical school when being admitted in the first lottery, and do not complete medical school when not being admitted in the first lottery. The latter implies that an applicant either does not reapply after losing the first lottery, reapplies one or more times and always loses, or wins a subsequent lottery but does not complete medical school. The first group of individuals who do not reapply after losing their first lottery seems to meet the concept of compliance closer than re-applicants who only comply because they also lose subsequent lotteries.

Therefore, we next define compliers in terms of the maximum number of times they are willing to participate in an admission lottery. "First-degree" compliers are individuals who complete medical school when being admitted in their first lottery and do not reapply when losing this lottery. "Second-degree" compliers reapply once when losing their first lottery and complete medical school when being admitted in one of these two lotteries. And similarly, we can define "third-degree" compliers and so on. The more often someone reapplies, the closer he or she is to an always taker. Below we show how we use the information about subsequent lotteries to separately identify the earnings returns to medical school for the different degrees of compliers.

We simplify the analysis to five types of applicants: first, second, third-degree compliers, never takers and always takers. Always takers are people who participate in at least four lotteries and complete medical school when being admitted. It is straightforward to extend the analysis to more groups of compliers, as will become clear below. Defining always takers as people who intend to participate in four or more lotteries does not affect the average returns for the other groups of compliers. Never takers are people who do not complete medical school irrespective of the result of the lotteries. By imposing monotonicity, we exclude defiers.

We first want to identify the shares of the five groups from our data. Let L_k be a dummy variable indicating whether a person would apply for the k^{th} lottery; 1 if yes, otherwise 0. LR_k is a dummy variable indicating whether a person won the k^{th} lottery (conditional on participating); 1 if yes, otherwise 0. C_N is a stochastic variable indicating whether the person is a never taker. A never taker is an applicant, who will not complete medical school when being admitted:

$$\Pr(C_N = 1) = \Pr(D = 0|LR_1 = 1)$$

Next, let C_k be a stochastic variable describing whether the person is a k^{th} degree complier. A first-degree complier does not reapply after losing the first lottery, and completes medical school when being admitted in the first lottery. The proportion of first-degree compliers can be identified by taking the proportion of losers of the first lottery that do not participate in a second lottery and correct for the fact that part of them will be never takers:

$$\Pr(C_1 = 1) = (1 - \Pr(C_N = 1))\Pr(L_2 = 0|LR_1 = 0)$$

Similarly, we consider the second-degree and third-degree compliers:

$$\Pr(C_2 = 1) = (1 - \Pr(C_N = 1))\Pr(L_3 = 0|LR_2 = 0, L_2 = 1)\Pr(L_2 = 1|LR_1 = 0),$$

and,

$$\Pr(C_3 = 1) = (1 - \Pr(C_N = 1))\Pr(L_4 = 0|LR_3 = 0, L_3 = 1)$$
$$\Pr(L_3 = 1|LR_1 = 0, LR_2 = 0)\Pr(L_2 = 1|LR_1 = 0).$$

Finally, let C_A be a stochastic variable indicating whether the person is an always taker. Given that we already derived expressions for the share of the four other groups, the probability of being an always taker is:

$$\Pr(C_A = 1) = 1 - \Pr(C_3 = 1) - \Pr(C_2 = 1) - \Pr(C_1 = 1) - \Pr(C_N = 1)$$

For tractability we impose that the probability of being a never taker is the same in every lottery, which is consistent with our data. In each lottery the fraction of admitted students that does not complete medical school is very similar.[24]

Next, we consider the probability of being a complier in our empirical analyses, which are individuals for which the outcome of the first lottery determines whether or not to complete medical school. This holds for first-degree compliers, and second-degree and third-degree compliers who lose all subsequent lotteries after the first lottery:

$$\Pr[D(LR_1 = 1) - D(LR_1 = 0) = 1] =$$
$$P(C_1 = 1) + P(C_2 = 1)P(LR_2 = 0) + P(C_3 = 1)P(LR_2 = 0)P(LR_3 = 0)$$

Using this characterization of the compliers in our estimations, we can then rewrite the right hand-side of equation (3) as a weighted average of the potential outcomes for the

[24] In the first lottery the fraction of never takers is 0.176. For the second and third lottery these fractions are 0.181 and 0.164, respectively.

different groups of compliers:

$$\frac{E[Y|LR_1=1]-E[Y|LR_1=0]}{P(D=1|LR_1=1)-P(D=1|LR_1=0)} = \\ (E[Y(1)-Y(0)|C_1=1]P[C_1=1] \\ + E[Y(1)-Y(0)|C_2=1]P[C_2=1]P[LR_2=0] \\ + E[Y(1)-Y(0)|C_3=1]P[C_3=1]P[LR_2=0]P[LR_3=0])/ \\ (P[C_1=1]+P[C_2=1]P[LR_2=0]+P[C_3=1]P[LR_2=0]P[LR_3=0]) \quad (4)$$

We now restrict the sample to people who lost their first lottery and applied for the second lottery. For these applicants, the result of the second lottery is random (conditional on the lottery category). Again using instrumental variables estimation, we estimate for the compliers to the result of the second lottery the earnings returns to medical school. This group of compliers to the result of the second lottery consists of second-degree and third-degree compliers. Recall that first-degree compliers never get to the second lottery.

The Wald estimate using applicants to the second lottery can be expressed as a weighted average of potential outcomes of second-degree and third-degree compliers:

$$\frac{E[Y|LR_2=1]-E[Y|LR_2=0]}{P(D=1|LR_2=1)-P(D=1|LR_2=0)} = \\ (E[Y(1)-Y(0)|C_2=1]P[C_2=1] \\ + E[Y(1)-Y(0)|C_3=1]P[C_3=1]P[LR_3=0])/ \\ (P[C_2=1]+P[C_3=1]P[LR_3=0]) \quad (5)$$

Repeating the instrumental variables estimation for applicants to the third lottery gives the earnings returns for compliers to the result of the third lottery. This group only consists of third-degree compliers, which gives the following expression:

$$\frac{E[Y|LR_3=1]-E[Y|LR_3=0]}{P(D=1|LR_3=1)-P(D=1|LR_3=0)} = E[Y(1)-Y(0)|C_3=1] \quad (6)$$

With the estimate of the earnings returns for third-degree compliers in equation (6) and

Table 6. Estimates of earnings returns to medical school for different lotteries and degrees of compliers, by year since first lottery participation

	Lottery 1		Lottery 2		Lottery 3		Compliers		
$t-\tau$	N	IV	N	IV	N	IV	1st degree	2nd degree	3rd degree
0	2159	−1.2 (0.3)***	879	0.1 (0.5)	280	0.2 (1.0)	-2.3 (0.7)***	0.0 (0.7)	0.2 (1.0)
1	4607	−2.5 (0.3)***	1898	−1.0 (0.4)**	674	−0.1 (0.8)	-3.7 (0.6)***	-1.3 (0.4)**	-0.1 (0.8)
2	7167	−0.7 (0.3)***	2995	−1.4 (0.3)***	1043	−1.2 (0.6)*	0.0 (0.4)	-1.5 (0.4)***	-1.2 (0.6)*
3	9885	0.1 (0.3)	4243	−0.6 (0.3)*	1500	−1.9 (0.5)***	0.7 (0.5)	-0.1 (0.4)	-1.9 (0.5)***
4	12,438	−1.3 (0.3)***	5374	−0.2 (0.4)	1911	0.2 (0.5)	-2.2 (0.5)***	-0.4 (0.5)	0.2 (0.5)
5	14,952	−7.1 (0.4)***	6371	−3.4 (0.5)***	2263	−1.2 (1.0)	-10.1 (0.6)***	-4.3 (0.7)***	-1.2 (1.0)
6	17,154	−11.0 (0.4)***	7245	−10.1 (0.5)***	2546	−5.1 (0.9)***	-11.7 (0.8)***	-12.0 (0.7)***	-5.1 (0.9)***
7	18,945	11.3 (0.8)***	7854	−13.7 (0.7)***	2731	−12.5 (0.9)***	31.6 (1.3)***	-14.2 (0.9)***	-12.5 (0.9)***
8	20,705	20.3 (0.9)***	8394	8.5 (1.5)***	2873	−14.8 (1.2)***	29.8 (2.1)***	17.5 (1.9)***	-14.8 (1.2)***
9	22,183	13.8 (0.7)***	8688	14.7 (1.1)***	2926	0.2 (1.7)	13.0 ()	20.3 ()	0.2 (1.7)
10	23,484	9.8 (0.7)***	8886	8.6 (1.0)***	2944	7.5 (1.8)***	10.7 (1.3)***	9.0 (1.3)***	7.5 (1.8)***
11	24,849	7.6 (0.7)***	9102	3.7 (0.9)***	2971	3.7 (1.7)**	10.7 (1.4)***	3.8 (1.3)***	3.7 (1.7)**
12	22,608	7.7 (0.8)***	8190	4.3 (1.1)***	2676	2.7 (1.7)*	10.4 ()	4.9 ()	2.7 (1.7)*
13	20,117	9.3 (0.9)***	7142	3.2 (1.3)**	2259	1.7 (1.8)	14.2 ()	3.8 ()	1.7 (1.8)
14	17,568	11.3 (1.3)***	6059	6.2 (1.8)***	1904	−0.8 (2.5)	15.5 ()	8.8 ()	-0.8 (2.5)
15	14,894	16.2 (1.7)***	4840	8.6 (2.5)***	1453	2.6 (4.7)	22.4 (3.5)***	10.9 (3.8)***	2.6 (4.7)
16	12,410	25.2 (2.2)***	3767	11.9 (3.3)***	1067	10.3 (6.7)	35.9 (4.6)***	12.6 (5.6)**	10.3 (6.7)
17	9949	29.4 (2.7)***	2803	20.7 (4.3)***	731	13.0 (8.2)	36.4 (6.1)***	23.6 (6.4)***	13.0 (8.2)
18	7808	30.6 (3.6)***	1962	18.9 (5.9)***	465	6.9 (11.5)	40.1 (7.8)***	23.5 (8.6)***	6.9 (11.5)
19	6056	42.4 (4.2)***	1363	19.3 (8.1)**					
20	4306	41.8 (5.3)***	829	8.2 (12.1)					
21	2784	42.3 (6.9)***	493	31.5 (15.4)**					
22	1430	39.4 (7.3)***							

Notes: Robust standard errors and bootstrapped standard errors (500 replications) for 1st degree and 2nd degree compliers in parentheses. The total number of individuals in lottery 1, 2 and 3 are respectively 25,491, 13,672 and 5,910. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Every cell in this table represents a separate regression, which include controls for gender, ethnicity, age in the first lottery year, and for each lottery the lottery category in that lottery , year of lottery and interaction terms of the year of lottery and lottery category.

estimates of the shares of second-degree and third-degree compliers, we can recursively compute the earnings returns for second-degree compliers from equation (5). Likewise, with estimates of the earnings returns of third-degree and second-degree compliers and of the shares of first-degree, second-degree and third-degree compliers, we recover the earnings returns for first-degree compliers from equation (4).

The first three columns in Table 6 report IV estimates of the earnings returns to medical school by year since first lottery using the result of the first, second and third lotteries as instrumental variables. We do not report results for years in which the sample size falls below 300 as than the number of observations becomes to small to estimate the model with all the controls. The results in the first column repeat those from column (4) in Table 3, the results in the second column are obtained when restricting the sample to people who lost the first lottery and apply to the second. Likewise the results in the third column are obtained when restricting the sample to people who lost the first and second lotteries and apply to the third.

The results in the third column pertain to third-degree compliers. Using the procedure outlined above, the next two columns present estimates of the earnings returns for first-degree and second-degree compliers.[25] For some years after the first lottery, the numbers of observations are rather small and the estimates imprecise, but for the years with enough observations the pattern seems clear. The earnings returns for first-degree compliers are larger than for second-degree compliers, which in turn is larger than for third-degree compliers. Figure 10 shows the returns for the different degrees of compliers. In order to correct for the fact that second and third-degree compliers study longer we now have years since last lottery on the X-axis.

As stated above, if individual's preferences for medical school and returns to medical school would be independent, we would expect returns to medical school to be higher for individuals who are willing to participate in more admission lotteries. The results show the opposite pattern, higher-degree compliers have lower earnings returns than the

[25]The shares of the different groups are: 0.229 for first-degree compliers, 0.230 for second-degree compliers, 0.154 for third-degree compliers, 0.210 for always takers, and 0.176 for never takers. The probability to lose the first lottery ($\Pr(LR_1 = 0)$) equals 0.536. For the second and third lotteries these probabilities are 0.582 and 0.574, respectively.

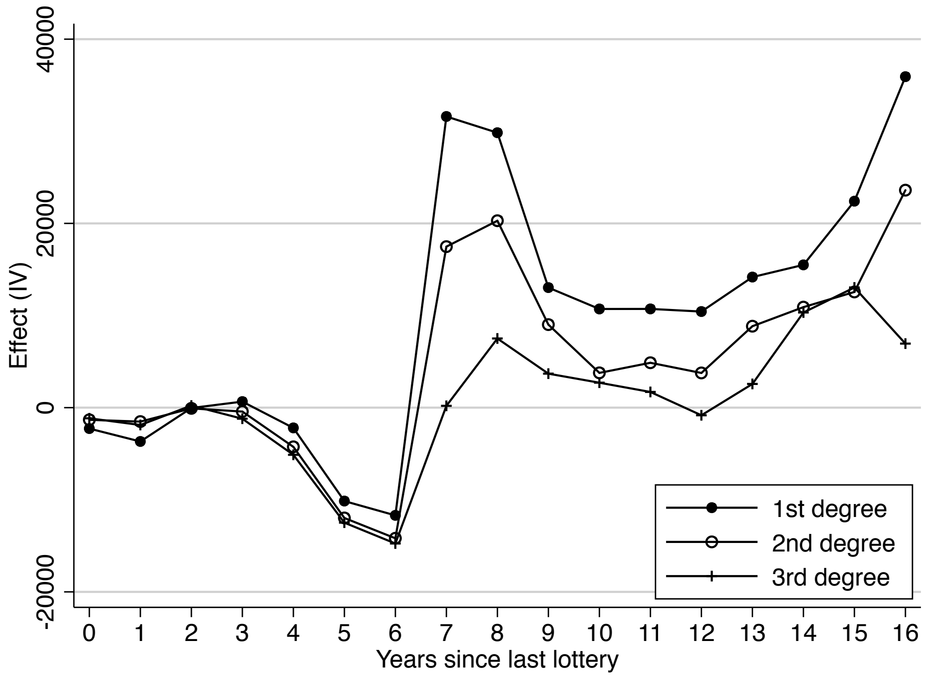


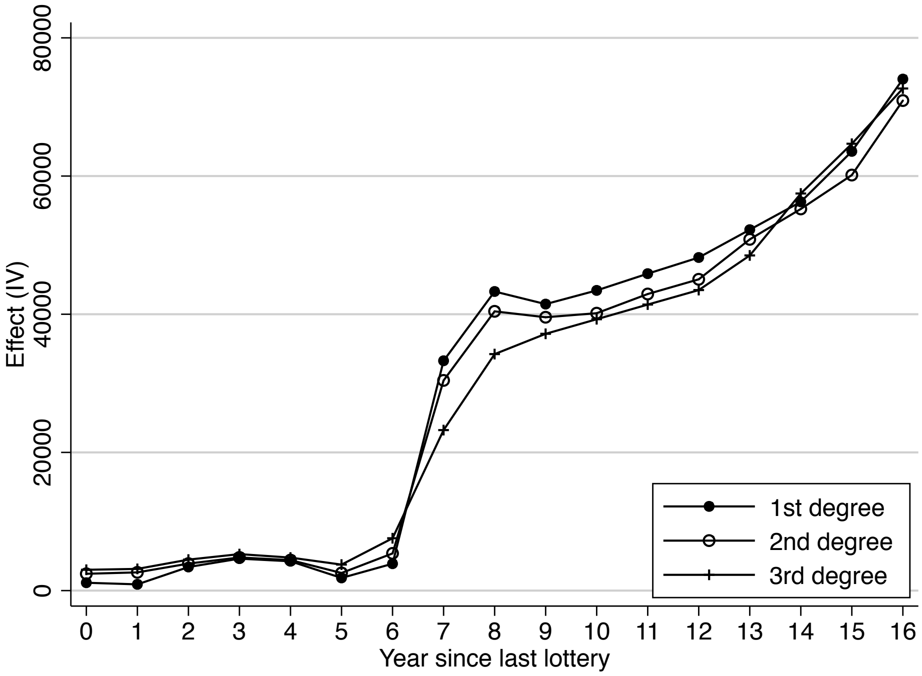
Figure 10. Estimates of earnings returns to medical school completion for different lotteries and degrees of compliers, by year since first lottery

lower-degree compliers. This can only be explained if preferences for medical school are negatively related to earnings returns.

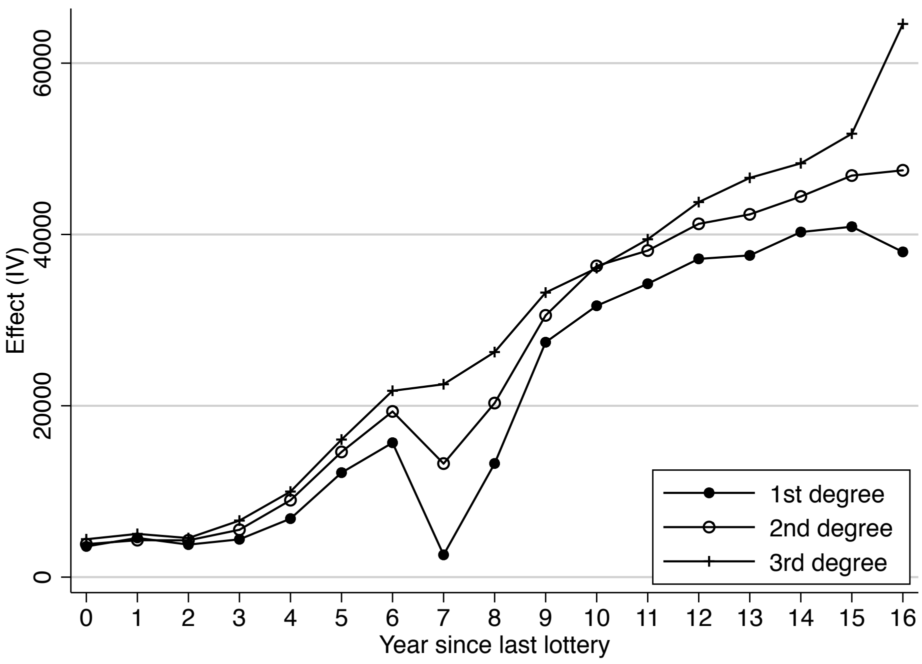
Figure 11 shows the predicted levels of earnings for different degrees of compliers. Earnings as a doctor are very similar for the different degrees of winning compliers, so the effect we find is mainly driven by the difference in earnings between degrees of losing compliers: Losing third-degree compliers earn more than losing first-degree compliers. A tentative conclusion could be that first-degree compliers are less motivated, both when it comes to reapplying for medical school as to their performance in their second best option. Third-degree compliers on the other hand, try harder to get in to medical school and are more motivated to make the best out of their second best option when they do not manage to get into medical school.

6.2 Validity of the instrument

We use the result of the first lottery as instrumental variable to estimate the returns to medical school. In Table 3 we have shown that this instrument has a strong impact on the endogenous variable. We argued that the exclusion restriction holds by virtue



(a) Completed



(b) Not completed

Figure 11. Predicted earnings for different degree compliers

of the randomization of the lottery. The exclusion restriction does not hold, however, when lottery losers are disappointed and, therefore, do worse, or when people become less motivated if they learn that they cannot follow the study of their first choice. This may lead to an overestimation of the earnings returns to medical school. To assess the importance of this mechanism, we analyze admission data from two other studies that have or had a quota and a centrally executed admission lottery: dentistry and international business studies. Just like medical school, the diploma of dentistry provides access to a highly regulated occupation. Completing dentistry might therefore also give access to monopoly rents. A diploma in international business studies does not provide access to a regulated labor market. We consequently assume that admission to this study does not generate a monopoly rent. If we still find an earnings difference between winning and losing compliers for this study, this could be the result of losers being disappointed and/or being less motivated for the study of their second choice.

Estimates of the earnings return to dentistry and international business studies are presented in Figure 12.[26] For dentistry the effect on earnings is positive and substantial from the fifth year since the first lottery onwards.[27] The effect is around 50,000 euro per year from the seventh to the thirteenth year and increases to 70,000 euro in the sixteenth year. These returns are higher than the returns to medical school in the same years, and show that also in the market for dentists substantial rents are present. The effect of completing international business studies on earnings is only significantly different from zero in one of the fourteen observation years. While we acknowledge that the level of disappointment from losing the lottery for medical school may differ from the level of disappointment from losing the lottery for international business studies, the estimates are still informative. They show that at least for international business studies there is no independent effect of losing the lottery on future earnings.

[26]The lottery for dentistry started later than the lottery for medical school, so we can only estimate the effect on earnings up to the sixteenth year after the first lottery. The lottery for international business studies was abolished in 1992 so for international business we can only estimate the effect from the seventh year on.

[27]In the years relevant for this paper the nominal duration of the dentistry study was five years.

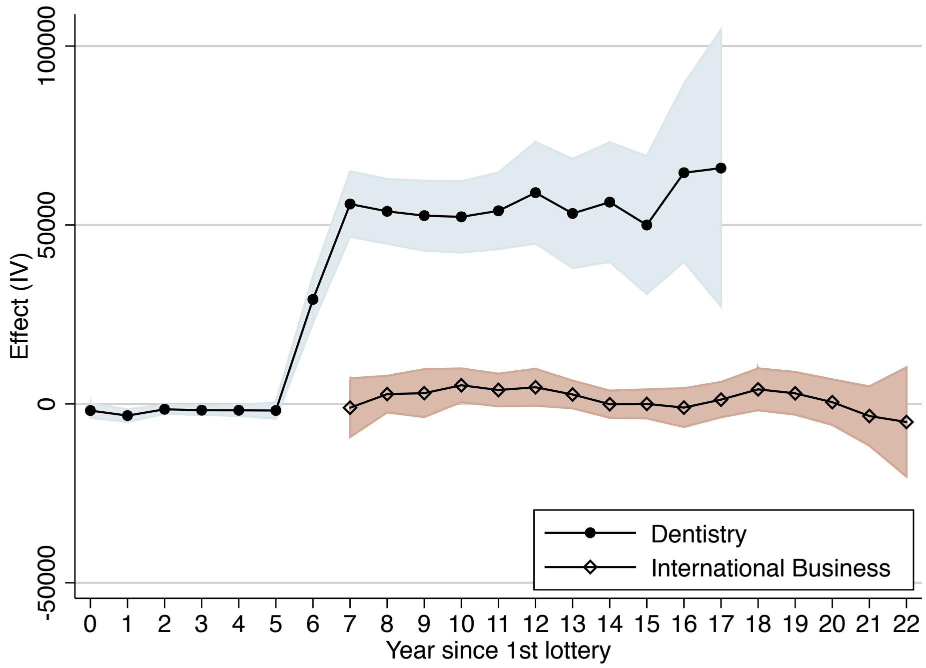


Figure 12. Estimated earnings returns to completing dentistry and international business studies $t - \tau$ years after first applying

7 Conclusion

Our empirical results provide evidence for substantial earnings returns to medical school. In each year after graduating these returns are at least 20 percent compared to the second-best study, and the returns increase to almost 50 percent 22 years after first applying to medical school. Only a small part of this earnings difference can be attributed to differences in working hours or more investment in human capital. If we interpret the remainder of the earnings returns as monopoly rent, this also explains why the number of applicants is substantially higher than the number of available slots in medical schools.

Releasing the quota might reduce the monopoly rents of doctors. If we assume that wages in the applicant's next-best option are not influenced by a release of the quota such a release can reduce doctors' earnings to the level in their next-best option.[28] Releasing the quota is costly in a situation in which the government heavily subsidizes study costs, as is currently the case in the Netherlands. The costs of attending medical school are

[28] Earnings levels in applicants' next-best option will be affected if releasing the quota significantly reduces labor supply in these sectors. In most alternative fields in which rejected medical school applicants apply they form only a small proportion of the total amount of students (for example law or psychology), so this is not likely to be the case.

much higher than the costs of other study programs. The total costs of attending medical school are estimated to be at least 167,000 euros compared to an average amount of 55,000 euros for other university study programs (Houkes-Hommes, 2009).[29] Students pay only a tuition fee of around 1000 euros per year, which is not differentiated across studies. Furthermore, the majority of the medical school students starts a specialization track. The costs of a specialization track are completely covered by the government and range from 40,000 to 145,000 euro.[30]

Releasing the quota may not only increase public expenditures on university education, but it is often argued that due to supplier-induced demand also health-care costs may increase. However, one might question if the coexistence of high private returns and high public investment is desirable. Policy makers can either consider maximizing earnings for individuals in the medical profession or shift part of the study costs to students. Under the conservative assumptions underlying Table 4, there is sufficient scope for medical school students to pay a larger share of their study costs. This might also allow the government to increase the number of available places without increasing public expenditures. At the same time higher costs can reduce the number of applicants for medical school. An increase in the supply of doctors and the resulting reduction of their earnings will also reduce the number of applicants.

References

Anderson, G. M., Halcoussis, D., Johnston, L., and Lowenberg, A. D. (2000). Regulatory barriers to entry in the healthcare industry: The case of alternative medicine. *Quarterly Review of Economics and Finance*, 40(4):485–502.

Burstein, P. L. and Cromwell, J. (1985). Relative incomes and rates of return for U.S. physicians. *Journal of Health Economics*, 4(1):63 – 78.

[29]Part of the difference in costs reflects the fact that medical school takes longer than the alternative study programs.

[30]The specialization tracks are an exception among other post-graduate programs; in most cases the government does not bear the (full) costs of post-graduate education.

Friedman, M. and Kuznets, S. (1954). *Income from Independent Professional Practice*. NBER Books. National Bureau of Economic Research, Inc.

Herfs, P. (2009). International medical graduates in the Netherlands. Phd-dissertation, Utrecht University.

Houkes-Hommes, A. (2009). De kosten van verruimen of loslaten van de numerus fixus. Report, SEO.

Hurley, J. (2000). An overview of the normative economics of the health sector. In Culyer, A. J. and Newhouse, J. P., editors, *Handbook of Health Economics*, volume 1A. Elsevier.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Imbens, G. W. and Rubin, D. B. (1997). Estimating outcome distributions for compliers in instrumental variables models. *Review of Economic Studies*, 64(4):555–574.

Kugler, A. D. and Sauer, R. M. (2005). Doctors without borders? Relicensing requirements and negative selection in the market for physicians. *Journal of Labor Economics*, 23(3):437–466.

OECD (2010). *Health at a glance 2010*. OECD, Paris.

RVZ (2010). Numerus fixus geneeskunde: Loslaten of vasthouden. Report, Raad van Gezondheid (RVZ).

Schafer, W., Kroneman, M., Boerma, W., van den Berg, M., Wester, G., Deville, W., and van Ginneken, E. (2010). The Netherlands: Health system review. *Health Systems in Transition*, 12(1):1–229.

Simoens, S. and Hurst, J. (2006). The supply of physician services in OECD countries. OECD Health Working Papers 21, OECD Publishing.

A Results Tables

Table A1. Fraction p admitted and number of applicants N by year and lottery category

	A		B		C		D		E		F		Total	
Year	p	N	p	N	p	N	p	N	p	N	p	N	p	N
1988	1.00	29	1.00	96	0.89	179	0.75	495	0.62	537	0.54	749	0.67	2085
1989	1.00	30	1.00	84	0.96	158	0.80	429	0.66	531	0.58	697	0.71	1929
1990	1.00	36	1.00	111	0.87	194	0.71	468	0.59	571	0.51	746	0.64	2126
1991	1.00	41	0.89	130	0.76	201	0.63	547	0.50	649	0.43	881	0.56	2449
1992	1.00	51	0.84	113	0.72	235	0.59	600	0.48	689	0.42	1036	0.53	2724
1993	0.93	44	0.72	167	0.62	241	0.51	702	0.41	847	0.36	1299	0.45	3300
1994	0.89	61	0.69	208	0.58	389	0.48	905	0.39	1034	0.33	1331	0.43	3928
1995	0.80	88	0.62	265	0.51	430	0.41	982	0.34	1024	0.31	1402	0.39	4191
1996	0.74	97	0.58	283	0.48	494	0.39	1084	0.32	1119	0.27	1496	0.36	4573
1997	0.72	117	0.54	310	0.45	498	0.37	1114	0.31	1129	0.26	1486	0.35	4654
1998	0.75	106	0.56	332	0.50	492	0.39	1121	0.32	1041	0.28	1325	0.37	4417
1999	1.00[a]	87	1.00[a]	341	0.52	421	0.42	1025	0.33	898	0.29	1146	0.43	3918
Total	0.86	787	0.73	2440	0.59	3932	0.49	9472	0.41	10,069	0.36	13,594	0.46	40,294

[a]In 1999 a reform was implemented which implied that from that year on applicants with a GPA above 8 (category A and B) are automatically admitted

Table A2. IV estimates of effects of medical school completion on earnings and hours, by year since first lottery and gender

	Earnings		Hours	
$t-\tau$	Men	Women	Men	Women
0	−1.7 (0.6)***	−0.9 (0.4)**		
1	−2.5 (0.5)***	−2.5 (0.3)***		
2	−0.6 (0.4)	−0.7 (0.3)**		
3	0.1 (0.5)	0.0 (0.3)		
4	−1.7 (0.5)***	−1.1 (0.3)***		
5	−5.4 (0.6)***	−8.4 (0.4)***		
6	−8.2 (0.7)***	−13.2 (0.5)***		
7	6.9 (1.1)***	14.9 (1.1)***	169 (160)	268 (128)**
8	14.7 (1.5)***	24.9 (1.1)***	277 (98)***	601 (76)***
9	11.2 (1.1)***	15.9 (0.9)***	146 (63)**	385 (48)***
10	8.4 (1.1)***	10.9 (0.8)***	199 (47)***	288 (35)***
11	5.7 (1.2)***	9.1 (0.9)***	88 (36)**	150 (31)***
12	5.1 (1.4)***	9.9 (0.9)***	122 (32)***	142 (30)***
13	7.3 (1.6)***	11.0 (1.1)***	105 (30)***	149 (31)***
14	8.9 (2.1)***	13.3 (1.5)***	112 (29)***	134 (33)***
15	15.7 (2.9)***	16.6 (1.9)***	114 (29)***	100 (35)***
16	30.4 (4.0)***	21.0 (2.4)***	62 (34)*	67 (36)*
17	37.5 (5.0)***	23.3 (2.9)***	71 (37)*	82 (38)**
18	35.8 (6.9)***	26.7 (3.7)***	50 (36)	85 (40)**
19	53.2 (7.9)***	34.9 (4.3)***	71 (40)*	90 (42)**
20	47.7 (10.2)***	37.0 (5.4)***	76 (48)	122 (51)**
21	42.0 (12.6)***	41.0 (7.4)***	−14 (52)	42 (62)
22	36.3 (12.7)***	40.0 (8.2)***	−35 (68)	151 (85)*

Notes: Standard errors in parentheses. Total number of individuals is 25,491, of which 10,661 are men and 14,880 are women. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Every cell in this table represents a separate regression, which include controls for ethnicity, age in the first lottery year, lottery category, year of first lottery and interaction terms of the year of first lottery and lottery category.

Table A3. IV estimates of effects of medical school completion on earnings, by year since first lottery and lottery category

$t-\tau$	B	C	D	E	F
0	−1.0 (1.7)	−0.9 (1.0)	−1.1 (0.6)**	−1.5 (0.6)***	−1.2 (0.8)
1	−12.1 (6.5)*	−1.1 (0.8)	−1.9 (0.5)***	−2.5 (0.5)***	−2.9 (0.5)***
2	0.5 (1.3)	1.3 (0.7)*	0.0 (0.5)	−0.9 (0.5)**	−1.8 (0.5)***
3	1.3 (1.7)	0.9 (0.7)	0.2 (0.5)	0.2 (0.4)	−0.6 (0.6)
4	−1.3 (1.2)	−0.7 (0.8)	−0.7 (0.5)	−1.9 (0.5)***	−1.6 (0.6)***
5	−9.2 (1.9)***	−7 (1.1)***	−7.4 (0.8)***	−6.8 (0.7)***	−6.8 (0.7)***
6	−11.3 (2.2)***	−9.4 (1.4)***	−10.7 (0.9)***	−11.3 (0.8)***	−11.3 (0.8)***
7	32.9 (5.5)***	27.4 (3.3)***	18.4 (1.6)***	7.9 (1.3)***	0.9 (1.2)
8	41.7 (5.9)***	31.0 (3.2)***	28.4 (1.7)***	17.8 (2.0)***	10.5 (1.4)***
9	19.7 (4.3)***	13.5 (2.3)***	18.1 (1.4)***	12.8 (1.3)***	10.8 (1.2)***
10	16.4 (4.2)***	7.0 (2.1)***	11.7 (1.3)***	8.6 (1.2)***	9.4 (1.2)***
11	17.5 (4.6)***	8.7 (2.4)***	8.1 (1.4)***	7.3 (1.4)***	6.3 (1.1)***
12	10.2 (5.4)*	9.3 (2.9)***	8.2 (1.6)***	6.9 (1.5)***	7.4 (1.3)***
13	7.1 (6.9)	12.7 (3.5)***	9.7 (2.0)***	8.2 (1.6)***	9.3 (1.5)***
14	−1.8 (10.7)	8.1 (4.6)*	9.9 (2.7)***	14.7 (2.3)***	11.5 (1.9)***
15	5.4 (14.7)	16.2 (7.2)**	16.8 (3.7)***	16.3 (3.0)***	16.6 (2.5)***
16	19.2 (17.6)	18.0 (10.7)*	26.6 (4.9)***	24.6 (3.8)***	26 (3.4)***
17	2.2 (31.7)	28.1 (11.0)***	37.8 (7.0)***	28.1 (4.7)***	27.4 (3.9)***
18	−125.3 (230.4)	35.5 (14.7)**	33.0 (9.2)***	28.1 (7.0)***	31.9 (4.8)***
19	74.3 (1937.3)	47.7 (14.0)***	40.4 (11.9)***	44.7 (7.2)***	41.9 (5.6)***
20	136.1 (40.2)***	66.1 (18.8)***	12.8 (18.2)	51.5 (8.7)***	44.2 (6.8)***
21	137.7 (57.2)**	72.8 (35.8)**	13.6 (26.2)	48.8 (11.0)***	44.4 (8.3)***
22	124.5 (56.9)**	72.3 (60.0)	46.5 (21.8)**	30.9 (13.4)**	40.6 (9.1)***

Notes: Standard errors in parentheses. Total number of individuals in categories B, C, D, E and F are respectively 1808, 2724, 6083, 6435 and 8441. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Every cell in this table represents a separate regression, which include controls for gender, ethnicity, age in the first lottery year and year of first lottery.

B Estimation of outcome distributions for compliers

Imbens and Rubin (1997) show how to derive distributions of outcome for both winning and losing compliers. Below we briefly review this approach. $Y_i(1)$ and $Y_i(0)$ denote the potential earnings with and without completing medical school, respectively. For each observation we observe the triple (LR_{1i}, D_i, Y_i). We cannot directly identify compliers from the data, but we can identify some never takers (for whom $LR_{1i} = 1$ and $D_i = 0$) and some always takers ($LR_{1i} = 0$ and $D_i = 1$). Because of the randomization, the instrument will be independent of a person's type, so in a large sample we can infer the distribution of $Y_i(1)$ for always takers and $Y_i(0)$ for never takers. These distributions are described by $g_a(y)$ and $g_n(y)$. Furthermore we know the population proportions ϕ_c, ϕ_a and ϕ_n of compliers, always takers and never takers, respectively.

The distributions of interest are the distributions of $Y_i(0)$ and $Y_i(1)$ for compliers, denoted as $g_{c0}(y)$ and $g_{c1}(y)$. These cannot be observed directly from the data because the group of lottery losers that do not complete medical school (with $LR_{1i} = 0$ and $D_i = 0$) consists of compliers and never takers. Analogously, in the outcome distribution of lottery winners that complete medical school ($LR_{1i} = 1$ and $D_i = 1$) there will be compliers and always takers.

We write the directly estimable distributions of Y_i for the subsample defined by $LR_{1i} = lr$ and $D_i = d$ as $f_{lr,d}(y)$. This implies that $g_a(y) = f_{01}(y)$ and $g_n(y) = f_{10}(y)$. Imbens and Rubin (1997) show that the distributions for the winning and losing compliers can be expressed in terms of the directly estimable distributions in the following way:

$$g_{c0}(y) = \frac{\phi_n + \phi_c}{\phi_c} f_{00}(y) - \frac{\phi_n}{\phi_c} f_{10}(y), \tag{A1}$$

and,

$$g_{c0}(y) = \frac{\phi_a + \phi_c}{\phi_c} f_{11}(y) - \frac{\phi_a}{\phi_c} f_{01}(y). \tag{A2}$$