The effect of receiving a Veni grant on long term research productivity

Layla Marie Ifzaren
Student ID: 457030

Erasmus University Rotterdam
Erasmus School of Economics

Supervisor: Josse Delfgaauw
Second assessor:

August 24, 2018

Contents

1 Introduction

The benefits derived from scientific research are typically both non-excludable and non-rivalrous. Therefore, economic theory predicts that scientific research would be undersupplied when left to the market. Hence, governments have been investing substantial resources in scientific research in order to solve for this market failure.

Governments can invest in scientific research in several ways. First and foremost, governments finance scientific research by providing universities with basic funding. Usually, money is simply divided amongst universities in lump-sum. The institutes then decide for themselves where the money will be invested. Another main funding instrument used by governments and science institutions in most OECD countries as well as many others outside the OECD is the competitive research grant. Unlike the basic public funding directly aimed at universities, competitive research grants are alloted to specific research projects that have a predetermined topic, budget and duration. These projects are selected based on a peer review assessment. In the Netherlands in 2014, the public financing of scientific research consisted for 30% of competitive research grants (van Dalen et al., 2015).

Because of their lengthy selection process, competitive grants require considerable time, money and effort to be invested by the allocating governmental institution as well as by candidates. The benefits from such a grant should therefore also be sizable in order for the use of the funding instrument to be justified.

The benefits from scientific research are difficult to measure, but are linked to and reasonably proxied by quantitative and qualitative research productivity. Therefore, the added value of a competitive grant could be explored by estimating its effect on research productivity.

In this paper I investigate the Veni grant, which is a competitive grant aimed at young researchers by the Netherlands' biggest provider of research grants, the Netherlands Organization for Scientific Research (NWO). I use a fuzzy regression discontinuity design in order to identify the causal effect of the Veni grant on research productivity. I thereby try to distinguish between short- and long-term productivity in order to separate direct from indirect grant effects.

I find that the Veni grant causes researchers to write more and higher quality publications *after* they finished their project funded by Veni grant money. This provides evidence that the Veni grant impacts the researcher's productivity rather indirectly than directly. I find that this impact can be partially or even completely explained by rejected grant applicants[1] dropping out of academia entirely. For the (potentially) remaining effect of the Veni grant, the most straightforward explanation is that awarded grant applicants can spend more money or time on acquiring knowledge and new skills. An alternative and

[1] In this paper I will refer to rejected grant applicants as losers or grant-losers. I will refer to awarded grant applicants as winners, Veni-winners or grant-winners.

more scientifically founded explanation, which *also* poses a reason for the higher amount of academic drop-outs among rejected grant applicants, is signaling, where winning a grant is a positive signal to the research community, which leads to more exposure and opportunities. Indeed, the signaling power of the Veni has been shown (Bol et al., 2018) to increase the chance of receiving a follow-up grant.

In further research it is important to disentangle between these explanations since the acquisition of knowledge and skills can be considered a justified cause for the grant, while signaling does not add to the productive potential of the researcher.

The rest of this paper is organized as follows. Section 2 reviews the previous literature. Section 3 describes the Veni grant and its allocation procedure. Section 4 explains the empirical strategy. Section 5 describes the data. In Section 6 the estimation results are presented. Section 7 provides a conclusion and discussion.

2 Literature review

In an ongoing attempt to find the most cost-effective method of research funding, the *university-level* comparison of competitive research grants with basic funding has received considerable attention in the literature.

Huffman and Evenson (2006) find that basic funding has a larger positive impact on agricultural research in the United States than competitive grant funding. More generally, Auranen and Nieminen (2010) try to determine whether more competitive research funding systems[2] increase efficiency in publication output. Their findings are inconclusive. These papers focus generally on comparisons between countries. Their identification strategies are not robust and therefore this literature delivers mostly correlational evidence.

Aghion et al. (2010) proxy university output by patents and the number of publications. With the use of instrumental variables they find that more competitive university funding systems increase university output.

Other research focuses on the basic- and competitive funding systems separately. When considered by itself, the relationship between competitive research grants and productivity is usually explored at the individual level instead of at the university level. Much of the concerning literature is interested in the predictive validity of the selection of grantees, meaning that the researchers aim to find out whether the grant-winners perform better than the losers, not whether this improved performance is caused by the grant.

The relationship between competitive grants and productivity has often been examined for a short-term period covering the funded research project. With no inference of causality, Campbell et al. (2010) find that the Canadian NCIC grant is positively related to later NCIC-supported research productivity. Similar correlational evidence indicating such a positive relationship is provided

[2]Which include all kinds of competition incentives besides the competitive grant funding.

by Hornbostel et al. (2009) who examine the German Emmy Noether Programme, by van Leeuwen et al. (2012) who consider three Dutch NWO grants, by Reinhard (2009) who scrutinizes the Swiss National Science Foundation and by Bornmann et al. (2010) who also consider two Dutch NWO grants. Bornmann et al. (2010) additionally find that the best-ranked rejected applicants are more productive than the awarded applicants.

The causal short-term impact of competitive grants on productivity has also been covered. Jacob and Lefgren (2011) find no effect of a grant from the National Institutes of Health (NIH) in the United States on research productivity in the first five years after the application.

The long term relationship between research grants and productivity has been explored as well. In line with their short-term results, Jacob and Lefgren (2011) also find no effect of a NIH grant on longer term productivity covering six to ten years after the application. Van den Besselaar and Sandström (2015) critically review literature on long term productivity of grantees (Saygitov, 2014; Bornmann, 2008; van den Besselaar, 2013; Decullier, 2014; Mutz, 2014; Gallo et al., 2014; Danthi, 2014; Kaltman, 2014). In the empirical part of this paper they use non-parametric tests to make a comparison between the whole of Veni-grant-winners and the 10% best performing losers, some nine years after the application. They find that the grantees were not significantly more productive in terms of publications and that the losers even performed better in terms of citation scores. The comparison of these two groups does give useful information about predictive validity, but it is not applicable to my question of causality because the 10% best performing losers were selected based on their high productivity after the grant decision. When comparing the winners to all of the losers, they find that productivity is significantly higher in almost all indicators used for the winners. Since there were no controls included in the analysis, these results are uninformative about the impact of the grant.

Specifically in the Netherlands, the question of causality has been addressed by Lanser and van Dalen (2013), Gerritsen, Plug and van der Wiel (2013) and most recently by Bol, de Vaan and van de Rijt (2018). On a micro-level, these Dutch papers make use of solid methods to estimate the causal effect of competitive grants. Lanser and van Dalen (2013) and Bol et al. (2018) find little or no effect of grants on researcher productivity.

In my thesis I largely repeat part of the work done by van den Besselaar and Sandström (2015) and Bol et al. (2018) by examining the Veni grant and its causal impact on researcher productivity in the long term. I aim to improve upon Besselaar and Sandströms work with the use of a larger dataset and a methodology that allows for the inference of causality. Bol et al. (2018) used a t-test to compare the productivity of Veni applicants whose proposals' scores had laid closely around the cutoff score. I try to emend their work by making use of a regression discontinuity design, which allows to make use of applications further away from the cutoff. Also, I make use of more sophisticated indicators, provided by the Centre for Science and Technology Studies, which are normalized across scientific fields and years.

3 The Veni grant

The Veni is extended by the Netherlands Organization for Scientific Research (NWO). The NWO is an independent institute with the legally established task to encourage quality and innovation in the sciences. It falls under the responsibility of the Ministry of Education, Culture and Science. In 2002, the NWO created the Veni in order to give young and talented researchers the opportunity to do research in their preferred field of interest. With this initiative, NWO tries to stimulate innovative research and intents to motivate young, talented individuals to pursue a career inside of academia.

The Veni is one of the three personal grants provided by the NWO under the Innovational Research Incentives Scheme (IRI). The other two grants are called the Vidi and the Vici. These three grants are intended for different phases in the scientific career of researchers. The Veni competition is open to researchers who obtained a PhD degree at most three years ago at the start of the year in which the competition takes place. The Vidi and the Vici grants are open to researchers who obtained a PhD degree at least three years ago at the start of the year in which the competition takes place.[3]

Since 2002 there have been one or two competition rounds for the Veni grant each year. In 2002, 2004 and 2006 there were two competition rounds, the other years only had one round each. Within each competition round, applicants need to hand in a research proposal to (not more than) one NWO-domain which represents a specific scientific field[4]. If the amount of proposals is four or more times as large as the available number of grants, the proposal will have to make it through a preselection. If the proposal is preselected, it is sent to at least two referees who are selected by staff members of the NWO as experts in that scientific field. These referees then individually and anonymously write a report in which they provide a qualitative judgment about the proposal. As a response, the applicant then has one week to write a rebuttal. Subsequently, the proposal, the reports and the rebuttal are sent to the selection committee. The domain-specific selection committee consists of people who are experts in that scientific domain and it is established by the NWO. The selection committee reads the reports and the proposal and assigns a score to the proposal, based on this information (I will refer to this score as the 'pre-interview score'). If this score is high enough, the applicant is invited for an interview. As a guideline, twice the amount of applicants should be invited as the number of available grants.[5] If the researcher is invited, his proposal is then given a priority score by the selection committee based on an assessment of the proposal, the reports and the interview (I also refer to the priority score as the 'post-interview score'). The allocation of grants depends on the priority score. A pre-specified number

[3]Technically, for Vidi applicants there is no strict rule that states they cannot apply when their promotion has taken place less than three years ago. They are nevertheless strongly advised by the NWO not to apply for this grant.

[4]The NWO-domains consist of: Earth and Life Sciences, Chemical Sciences, Physical Sciences, Health Research, Physics, Technical Sciences, Humanities and Social Sciences.

[5]Therefore, the applicants chances of receiving an invitation for an interview are largely dependent on the available grant funding and the total amount of applications.

of grants are available, which are allocated to the applicants that receive the best priority scores. It is worth emphasizing that the priority score is independent from the pre-interview score. It is not uncommon for the pre-interview scores to be higher in absolute terms than the priority scores.

4 Empirical strategy

In order to determine the effect of receiving a Veni grant on the research productivity of a researcher I conduct a quantitative analysis. I use a regression discontinuity design (Lee & Lemieux, 2010) for the identification of the causal effect.

Theoretically, a model that captures the relationship of interest between the grant and long term productivity can be described as follows:

$$Y_{ijt+\tau} = \alpha_0 + \alpha_1 * G_{ijrt} + \alpha_2 * X_{ijrt} + u_{ijrt}$$

Where $Y_{ijt+\tau}$ stands for productivity in the period covering year t to year τ, G_{ijrt} is a dummy variable indicating whether the applicant received the grant, X_{ijrt} is a vector of control variables and u_{ijrt} represents the error term containing all other determinants of productivity. α_1 is the effect of the grant on productivity and α_2 represents a vector of the effects of the control variables. Subscripts i,j,r and t stand for person ID, scientific field, round and year respectively.

A simple OLS regression of $Y_{ijt+\tau}$ on G_{ijrt} will likely yield a biased estimate of α_1, since there are likely omitted variables in u_{ijrt} which influence both the probability of receiving the grant and the researchers productivity in his following career. These variables might for example include talent, intelligence or appearance.

4.1 Regression discontinuity design

Fortunately, the grants are allocated according to a specific rule, which is ultimately based on the priority scores (s_{ijrt}) that the applicants receive. Based on the priority scores, applicants are assigned a rank (r_{ijrt}). The applicant with the best priority score is ranked number one. The applicant with the next best priority score is then ranked number two, and so on. The number of available grants in the competition determines the cut-off rank. According to the allocation rule, the receipt of the grant is determined by whether the rank is above or below the cutoff rank level (r^*_{jrt}).[6] I exploit this rule to identify a causal relationship between grant allotment and productivity by comparing the researchers who are a little above the cutoff (the treatment group) to those who are slightly below the cutoff (the control group).

[6]The cutoff rank level is defined as the average of the ranks of the worst ranking winner and the best ranking loser of the competition.

Ideally, I would have used the priority score instead of the rank as the running variable since the priority score would have provided information on the qualitative distance between the proposals[7]. However, because the priority scores in my dataset turned out to be a mixture of pre- and post-interview scores, G_{ijrt} is not perfectly determined by s_{ijrt}. The rank of the proposal (r_{ijrt}), on the other hand, is not a mixture of pre-interview and post-interview ranks. Therefore I decided to use r_{ijrt} instead of s_{ijrt} as the running variable in my analysis. Nevertheless, it still contains some random noise, meaning that there are some proposals with a rank lower than r^*_{jrt} which did not receive the grant and vice versa. Figure 1 visualizes this by showing the probability that an applicant received the Veni grant as a function of the normalized rank he received in his first application round.

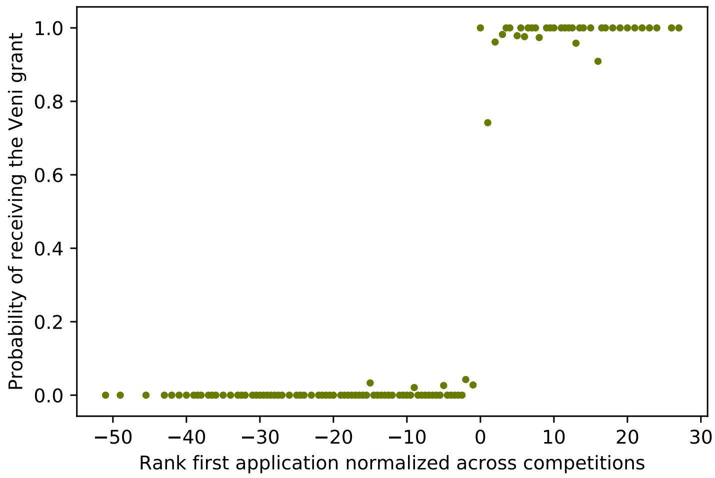

Figure 1: Probability of receiving the grant

As is visible from figure 1, either the allocation mechanism is not completely deterministic or there have been a few coding mistakes. Regardless of what caused the noise, a regular regression discontinuity (RD) analysis will now fail to identify the causal effect of the grant because the extension of the grant is now endogenous, even if the rank is controlled for. Basically, some of the applicants do not comply with the allocation rule either by not receiving the grant while their rank is lower than the cutoff rank or by receiving the grant while their rank is higher than the cutoff rank. Since the rank is no longer completely deterministic of the grant, a dummy variable indicating whether the rank is above or below the cutoff rank is needed as an instrumental variable to estimate the local average treatment effect (LATE) of the grant for the group of applicants who *do* comply to the allocation rule. When such an instrument is used, the

[7] A proposal with one rank higher could have received a marginally better or significantly better priority score.

RD design becomes a *fuzzy* RD design. The associated Two-stage-least-squares (2sls) estimation consists of two stages.

First stage:

$$G_{ijrt} = \alpha_0 + \alpha_1 * d_{ijrt} + f(r_{ijrt}) + \alpha_2 * X_{ijrt} + e_{ijrt}$$

Second stage:

$$Y_{ijt+\tau} = \beta_0 + \beta_1 * \hat{G}_{ijrt} + f(r_{ijrt}) + \beta_2 * X_{ijrt} + u_{ijrt}$$

In the first stage, G_{ijrt} is instrumented by a dummy ($d_{ijrt} = 1(r_{ijrt} > r^*_{jrt})$) which indicates whether the proposals rank was above or below the cutoff of that competition. In the second stage productivity is then regressed on $\hat{G}_{ijrt}$. Both in the first and second stage regressions a smooth polynomial of the rank in that competition ($f(r_{ijrt})$) which captures other (continuous) differences between individuals around the cut-off rank and a vector of control variables is included. I discuss the motivation behind the use of a smooth polynomial further below in this section.

One complication in the analysis is caused by reapplications. Researchers are allowed to apply for the same grant several times in different years as well as within a year with different proposals in multiple rounds; 12% of the candidates actually did so. Therefore, applicants who did not receive the grant in a certain round could receive it in a subsequent round of the same or another year. These candidates did not receive the grant in that moment, yet they could receive the grant at some point in their lifetime. If I do not control for this in my analysis, I am very likely to underestimate the impact of the grant on future productivity. In order to deal with this complication I limit my analysis to first time applicants[8] and estimate the effect of having ever received a Veni grant (either in this or in subsequent rounds). This implies that the independent variable G_{ijrt}, which takes value 1 if the applicant has received a particular Veni grant in round r and year t, will be replaced by G_{ij}, which takes value 1 if the applicant has received a Veni grant anytime ever. In the first stage, G_{ij} is then instrumented by a dummy ($d_{ij} = 1(r_{ij} > r^*_j)$) which indicates whether the proposal was above or below the rank in the first application round. In the second stage productivity is then regressed on $\hat{G}_{ij}$.

First stage:

$$G_{ij} = \alpha_0 + \alpha_1 * d_{ij} + f(r_{ij}) + \alpha_2 * X_{ij} + e_{ij}$$

Second stage:

$$Y_{ijt+\tau} = \beta_0 + \beta_1 * \hat{G}_{ij} + f(r_{ij}) + \beta_2 * X_{ij} + u_{ij}$$

For the unbiasedness of my regression discontinuity analysis, my assumptions about the relationship between the rank and the productivity indicators need to

[8] Technically, I limit my analysis to the first application for which a rank was available in my dataset (see Section 5). As a consequence, I use a second, third or fourth application for 189 applicants.

be of great accuracy. It is for this reason that I include a smooth polynomial of the rank into my main analysis where I use all the available data. The smooth polynomial allows for this relationship between rank and productivity, which can also be referred to as the functional form, to be anywhere between linear and cubic.

In the Appendix I repeat the fuzzy regression discontinuity analysis using smaller bandwidths including 75%, 50% and 25% of the data around the cutoff. The advantage of using a smaller bandwidth is that the two groups on either sides of the cutoff become more comparable. Also, the closer to the cutoff we look, the more linear the functional form becomes. As a result, estimates are less likely to be biased when a first degree polynomial of the rank is included. The disadvantage however, is that there may be too few observations for the estimation to produce statistically significant results.

As is visible from figure 2, the use of first time applications as an instrument reduces the strength of the first stage. The validity of the fuzzy design could potentially be threatened by the weak instruments problem (Staiger & Stock, 1997). Staiger and Stock (1997) suggest that when the instrument has an F-value higher than ten, it can be considered strong enough. In Table 7 of Section 6 I show that the first stage relationship between the dummy indicating whether the rank was above or below the cutoff in the first application round and ever receiving the grant is very strong.

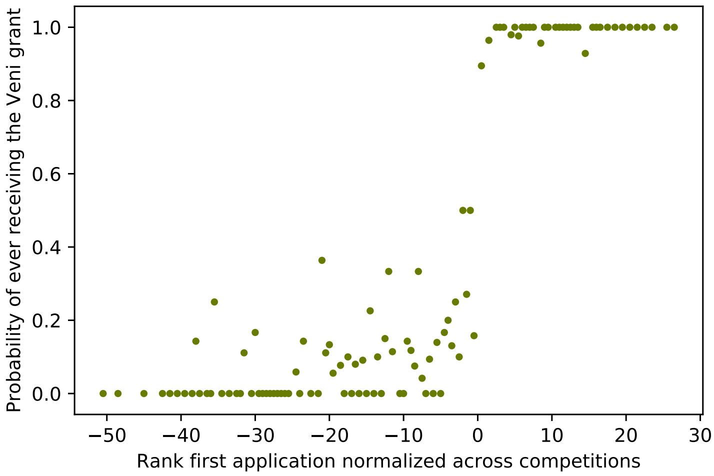

Figure 2: Probability of ever receiving the grant

Another potential concern would be that the effect of the Veni could be confounded by the effects of other competitive grants. This concern can be put to rest for the following reasons. First of all, other IRI grants could not confound the estimated effect of the Veni since the application requirements described previously exclude the possibility that the participants applied for the Vidi or

Vici while applying for the Veni. Second of all, the NWO is the biggest provider of competitive grants in the Netherlands, leaving only a few small grants with limited funds to confound the estimates in an trivial way. Third of all, outside of the Netherlands, the ERC Starting Grant extended by the European Research Council (ERC) would have been the largest and most viable alternative for rejected Veni applicants. This grant has a yearly competition round open to researchers with 2-7 years of experience since the completion of their PhD. Yet the chances of the estimates being confounded by this grant are low since the Starting Grant only began in the year 2007.

5 Data

I collected my data through desk research. In Section 5.1 I will describe the input data, which contains the information on grant proposals, their awards and scores. In Section 5.2 I will describe the output data, which contains the information on research productivity.

5.1 Input data

The input data was largely available in the CPB database from an earlier project called 'Up or out? How individual research grants affect academic careers in the Netherlands'. The paper was written by Gerritsen, Plug & van der Wiel, who received their data from the NWO in 2013. It contains data on research proposals for the Veni grant from postdocs from 2002-2010. It contains information on:

1. Whether the proposal was accepted or rejected for the IRI grant[9]
2. The priority score the proposal had received[10]
3. Whether the proposal had made it to the interview stage
4. The ranking of the proposal
5. The application date of the proposal

[9]As an additional check on this variable, I verified whether the proposal was accepted or rejected for the IRI grant using information from the NWO website about Veni-winners per round. For all winners in my dataset I checked whether they were present in the list of winners on the NWO website. As a result, I changed the status from accepted to rejected for four proposals, and I changed the status from rejected to accepted for two proposals. I dropped one duplicate observation that wrongly stated that the proposal had won the grant.

[10]Inconveniently, I found that the priority score contains both pre- and post-interview scores. I therefore needed to distinguish between the two types of scores using the *interview*-variable. However, for 920 out of 2,064 priority score observations this is not possible because they have missing values for *interview*. Thanks to the information that ís available, I know that 420 priority scores are pre-interview. This leaves 724 observations useful for an analysis using the priority score as a running variable and it led me to the decision to use the rank as the running variable instead.

6. The grant date of the proposal, which indicates when the applicant was informed about whether he or she received the grant
7. The research area the proposal had been in
8. Last name
9. Initials
10. Surname prefix(es)
11. Gender
12. Birth date

The complete dataset contains 6,076 observations. The ranks are missing for the years 2009 and 2010. My analysis is therefore limited to the years 2002-2008. With the missing values for 2009 and 2010 included, a total of 4,115 observations are missing for the ranking of the proposal.

I traced back the rounds to which the applications belonged by using information on the Veni-winners per round provided by the NWO website to separate the winners in my dataset into their respective rounds. Subsequently I categorized the losing applications into rounds using the application dates from the winning applications. If the application date from a losing application fell between the application dates of the first and the last winning application date from a particular round, then the losing application was considered to have taken place in that round. After using this method, 191 from 6,076 proposals still had information missing on the round they had been in because their application dates lie before the earliest winning proposal or after the latest winning proposal. 20 out of 1,961 proposals with information on rank had missing information on the round, which amounted to a 1 percent loss of data. At this point there were 1,942 proposals for which there were observations on rank, round and domain.

For 395 proposals, the rank variable does not contain a single rank, but rather a bandwidth wherein the proposal's rank falls. For all these proposals, I changed the bandwidth into it's average. Partly as a result of this procedure, I had for 523 proposals a rank that was shared by another proposal in the same competition, which in itself is no problem. I checked however whether there were applicants within a competition who had the same rank, while one received the grant and the other did not. I found that this was the case for 10 pairs of applications. For these application pairs, I changed the rank to missing for the proposal that had a different grant winning status from the proposals ranking one rank higher and lower since this was obviously the application that in reality must have had a different rank from the one written in the data.

According to the NWO rules people were not allowed to apply more than once every competition round. Still I observed 176 duplicate pairs[11] in my dataset. 7 of these pairs had one observation before the interview-phase and

[11] A 'pair' could also contain three observations

one observation after the interview-phase. For these pairs I deleted the duplicate observation from before the interview. 18 pairs turned out to be perfect duplicates, of these I removed whichever duplicate. For the last 114 pairs there was no information on *interview* available. Therefore, I removed the duplicate observation that had the worst rank, since that should have been the pre-interview observation.[12] Now 1,801 observations were left that contained a rank, a round and a domain.

I needed to determine the cutoff rank level, r^*_{jrt}, for each competition such that I could create the dummy instrument ($d_{ijrt} = 1(r_{ijrt} > r^*_{jrt})$).[13] I defined the cutoff as the number that equaled the total number of grant winners in that competition.

For the subsequent analysis I normalized the rank such that the proposals from different rounds and scientific domains could be analyzed all at once. The rank was simply normalized by subtracting the cutoff rank level from it within every competition. I then added 0.5 to all the ranks such that the worst ranking grant-winner had a rank of 0.5 and the best ranking grant-loser had a rank of -0.5.

5.2 Output data

I collected the output data myself from the Web of Science using an author disambiguation tool provided to me by the Centre for Science and Technology Studies (CWTS) in Leiden.[14] I collected the data per individual per year, starting from 2000 until 2016. This gives an overview of the development of an individual's research productivity over time, starting from *at least* two years before the Veni-application until *at least* eight years after.[15] I chose to estimate the effect of the grant on five productivity indicators:

1. Publications
2. The mean normalized citation score
3. The mean normalized journal score
4. The total normalized citation score
5. The total normalized journal score

Research productivity is defined by both quantity and quality which can be measured in several ways. I decided to estimate the effect of the Veni grant on both types of productivity separately, seeing as these effects could differ. Also, when trying to encapsulate quality and quantity in a productivity measure, it is not

[12] If one of the duplicates had a missing value for rank, it was discarded.

[13] For the imputation of the cutoff I used all applications, not just those that were in my estimation sample.

[14] I owe a great many thanks to Ed Noyons, Clara Calero Medina and the CWTS for providing me with their tool, their services, their help and their good advice.

[15] For some researchers I have data down till eight years before their application while for others I have data up till sixteen years after their application.

clear how much weight should be given to either of the two. I therefore decided to, first and foremost, estimate the grant effect on separate quantity and quality indicators. I used the number of publications as a straightforward quantitative measure of productivity. The mean normalized citation score (MNCS) and the mean normalized journal score (MNJS) were used as proxies for production quality. I deemed the MNCS to be a good indicator of research quality since its focus on mean citations makes sure that no weight is given to the *amount* of publications published. The same argument is valid for the MNJS, which is an indicator of research quality comparable to the impactfactor.

Since I also wanted to estimate a more general effect of the grant on productivity, I used the total normalized citation score (TNCS) and the total normalized journal score (TNJS) as well. These indicators increase with citations per publication/average journal publication as well as with the number of publications published. The weights that are given to publications and to citations in these indicators are arbitrary and do not reflect how I (or anyone ought to) think about their relative importance.

The normalized indicators are normalized for differences in citation practices between scientific fields. For the purpose of this field normalization, 4047 fields are distinguished. These fields are defined at the level of individual publications. Using a computer algorithm, each publication in the Web of Science is assigned to a field based on its citation relations with other publications. More elaborate information on the field assignment can be found at the website of CWTS Leiden Ranking.

The CWTS provided the productivity indicators per year. For each productivity indicator I constructed its total from its yearly values over three separate periods for every individual in my dataset. The first total indicator is used as a control variable and thus belongs to the input data. The other two total indicators are used as outcome variables. They are constructed to proxy short- and long-term research productivity respectively.:

1. The productivity indicator from 2-0 years before the application sums the yearly indicator over a two-year period before the applicant received the Veni.[16] I use this indicator as a control variable in my analysis since it gives information about the quality of the researcher before he received a grant.

2. The productivity indicator from 0-4 years after the application sums the yearly indicator for a four-year period from zero to four years after the individual received the Veni.[17] I will also refer to this indicator as short-term productivity. Most of it is directly funded with the Veni grant since the Veni grant funds research for a period of three years. The fourth year is added to this period to take into account the lag with which the grant-funded research is published.

[16] Or, in case the grant was never received, the yearly indicator is summed over the two-year period before the applicant first applied.

[17] Or, in case the grant was never received, the period starts at the first application year.

3. The productivity indicator from 5-8 years after the application sums the yearly indicator for a four-year period from the fifth up to and including the eighth year after the individual received the Veni. I use this indicator as a proxy for the researcher's long-term productivity. Presumably, the productivity measured by this indicator reflects post-Veni, alternatively funded research.

For the MNCS and the MNJS I calculated the yearly average of these three periods such that the regression coefficients can still be interpreted the way described later in this section.

The CWTS data collection method does not ensure the collection of all publications of all authors. Since I had less information on personal details for some applicants than I had for others, it is possible that the CWTS tool retrieved fewer publications from the Web of Science for these applicants. Especially the e-mail address was an important piece of information used by the CWTS disambiguation tool for the identification of a researcher's publications. Therefore, in my estimation sample I compared the percentage of grant losers for which I had an e-mail address with the percentage of grant winners for which one was available using a t-test shown below.

Table 1: Availability of e-mail address by grant winners and losers

	count(losers)	count(winners)	mean(lo)	mean(win)	p-value
email	552	569	0.89	0.97	0.00000

The e-mail address was missing significantly more often for grant-losers than for grant-winners. As a direct consequence, the CWTS tool had retrieved fewer of the losers publications from the WoS. Therefore, I included a dummy indicating whether I had the individuals e-mail address in every regression in order to prevent an overestimation of the effect of the grant on productivity.

In order to enable a proper interpretation of the regression results, I will explain more thoroughly how the MNCS, the MNJS, the TNCS and the TNJS are constructed in Sections 5.2.1, 5.2.2, 5.2.3 and 5.2.4. In the Appendix an elaborate explanation can be found on the data collection method.

5.2.1 Mean normalized citation score

Formally, the MNCS is represented as:

$$MNCS = \frac{1}{n}\sum_{1}^{n}\frac{c_i}{e_i}$$

Where c_i is the number of citations of publication i within a pre-specified timeframe and for a specific author, self citations excluded. It is divided by e_i, the expected number of citations of publication i given the year and the field in which it was published. The resulting normalized citation score (NCS) shows how much publication i had been cited compared to the average citation score of its field and year. Then the average of this NCS is taken for all publications of the author. The resulting MNCS should be interpreted as how much the author's publications were cited on average compared to how much an average publication in that year and field was cited. Hence, an MNCS of 2 would mean that the author's publications were cited twice as much as usual, an MNCS of 0.5 would mean that the author's publications were cited only half the usual amount.

5.2.2 Mean normalized journal score

The MNJS is written formally:

$$MNJS = \frac{1}{n}\sum_{1}^{n}\frac{mncs_i}{e_i}$$

Where $mncs_i$ is the average number of citations of all publications published in the same journal and the same year in which publication i was published. It is divided by e_i, which is the expected number of citations of *all* publications published in the same year and field as publication i. The resulting normalized journal score (NJS) shows how much an average publication from the journal of publication i had been cited compared to the average publication from publication i's field and year.Then the average of this NJS is taken for all publications of the author. The resulting MNJS should be interpreted as how much the journal's publications were cited on average compared to how much an average publication in that year and field was cited. Hence, an MNJS of 2 would mean that the journal's publications were cited twice as much as usual, an MNCS of 0.5 would mean that the journal's publications were cited only half the usual amount.

5.2.3 Total normalized citation score

Formally, the TNCS is written:

$$TNCS = \sum_{1}^{n}\frac{c_i}{e_i}$$

The TNCS is the sum of the normalized citation scores for all the author's publications in a given time period. Its value is harder to interpret than that of the MNCS because it can increase with the average number of citations per publication as well as with the total number of publications.

5.2.4 Total normalized journal score

The TNJS is written formally:

$$TNJS = \sum_{1}^{n}\frac{mncs_i}{e_i}$$

The TNJS is the sum of the normalized journal scores for all the author's publications in a given time period. Not unlike the TNCS, the TNJS is hard to interpret because it can increase with the average number of citations per journal publication as well as with the total number of publications from the author.

5.3 Estimation sample

The resulting output data was then combined with the input data. From the 1,801 proposals containing a rank and round, 1,599 proposals were first time applications, meaning that they were the first proposal the individual wrote in the competition for a Veni grant. The other 202 applications were re-applications from some of the applicants whose proposals had been rejected the first time. Hence, my data contained 1,801 applications from 1,599 applicants. After the data collection procedure, information on the productivity indicators was missing for 452 of the applicants[18], leaving 1147 applications useful for the regression discontinuity analysis. 32 Applicants received the Veni-grant after the year 2008. I dropped these applicants from my dataset to enable my regressions to be run using productivity indicators measured over a time-span of eight instead of six years. Without these applicants, the resulting estimation sample contained a total of 1,121 observations.

Section 5.3.1 presents the descriptive statistics. With the use of a balancing test in Section 5.3.2 I check whether grant winners are different from grant losers in observable characteristics. Finally, I compare my estimation sample to the original dataset in Section 5.3.3.

5.3.1 Descriptive statistics

Table 2 displays the descriptive statistics for the estimation sample. The descriptive statistics are shown for the whole estimation sample and for the group that ever won a Veni and the group that never won a Veni separately.[19]

The estimation sample contains (almost) equal observations on winners and losers. The mean normalized rank is higher for winners than for losers.[20] The min/max values for the normalized rank indicate that there are some proposals with a rank below the cutoff amongst the winners as well as some proposals with a rank above the cutoff amongst the losers. This shows that there is noise in the running variable. Both before and after their Veni application, grant winners were on average more productive in terms of publications, their

[18]The missing values indicated that the tool had not looked for the publications of these individuals on the WoS. Therefore, I could not know about these people whether they had or had not published any research during the eight years after their grant application. For some researchers, however, the CWTS tool found no publications on the WoS. A total of zero observations was found for 66 applicants in the short-term period and 108 applicants in the long-term period. These values were included into the analysis.

[19]For 'From the Netherlands' and 'Birth year' less observations are available. 1109 and 1004 for the full sample, 566 and 517 for the grant winners and 543 and 487 for the grant losers.

[20]The normalized rank is 'better' when it is greater than zero.

mean normalized citation score, their mean normalized journal score, their total normalized citation score and their total normalized journal score.

Table 2: Descriptive Statistics

	Full sample			Grant winners			Grant losers		
	mean	min	max	mean	min	max	mean	min	max
Ever received the grant	0.51	0.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
Normalized rank	-4.29	-48.50	25.50	4.40	-38.00	25.50	-13.24	-48.50	14.50
From the Netherlands	0.91	0.00	1.00	0.94	0.00	1.00	0.88	0.00	1.00
Birth year	1972.85	1946.00	1982.00	1973.17	1958.00	1982.00	1972.52	1946.00	1981.00
Female	0.34	0.00	1.00	0.39	0.00	1.00	0.30	0.00	1.00
0-4 years after application									
Publications	10.57	0.00	228.00	11.91	0.00	137.00	9.18	0.00	228.00
MNCS	1.35	0.00	10.70	1.51	0.00	8.90	1.19	0.00	10.70
MNJS	1.25	0.00	6.86	1.39	0.00	6.86	1.11	0.00	5.34
TNCS	16.98	0.00	307.20	19.61	0.00	284.04	14.27	0.00	307.20
TNJS	15.44	0.00	342.21	17.70	0.00	186.52	13.11	0.00	342.21
5-8 years after application									
Publications	15.00	0.00	238.00	17.65	0.00	238.00	12.26	0.00	195.00
MNCS	1.43	0.00	19.37	1.58	0.00	8.94	1.27	0.00	19.37
MNJS	1.30	0.00	17.87	1.44	0.00	7.72	1.16	0.00	17.87
TNCS	25.71	0.00	355.90	29.51	0.00	355.90	21.80	0.00	253.79
TNJS	23.43	0.00	321.45	26.94	0.00	321.45	19.82	0.00	263.01
2-0 years before application									
Publications	5.02	0.00	111.00	5.53	0.00	61.00	4.49	0.00	111.00
MNCS	1.51	0.00	18.20	1.73	0.00	18.20	1.28	0.00	11.54
MNJS	1.44	0.00	12.93	1.66	0.00	12.93	1.22	0.00	8.52
TNCS	7.75	0.00	142.25	9.15	0.00	142.25	6.31	0.00	120.13
TNJS	7.28	0.00	171.62	8.59	0.00	83.79	5.92	0.00	171.62
N		1121			569			552	

5.3.2 Balancing test

For the same estimation sample I check whether the winners and losers are significantly different in observable characteristics with the use of a t-test displayed in Table 3.[21]

Table 3: Balancing test

	Grant losers = (1) Grant winners = (2)		
	mean(1)	mean(2)	p-value
From the Netherlands	0.88	0.94	0.00015
Birth year	1972.52	1973.17	0.00955
Female	0.30	0.39	0.00196
Publications 2-0 years before application	4.49	5.53	0.00237
MNCS 2-0 years before application	1.28	1.73	0.00000
MNJS 2-0 years before application	1.22	1.66	0.00000
N	552	569	

Grant winners are significantly more likely to be female and from the Netherlands. Compared to grant losers they are publishing more, they do so in higher quality journals and these publications get cited more in the two years before they apply for the grant. These differences between the losers and winners should not impose any problems on the estimation of the effect of the grant on productivity as long as they are not discontinuous at the cutoff.

In order to check for this discontinuity, I regress[22] these observable characteristics on the dummy indicating whether the normalized rank is above or below the cutoff level. In this regression I control for the normalized rank itself, round, domain, e-mail and the other observable characteristics. The estimated effects of being above the cutoff are shown in Table 4. In line with the findings of Bol et al. I find that applications that rank above the cutoff level are significantly more likely to be from women. I also find that grant winners are more likely to be younger and that they publish in higher quality journals before their Veni application.

[21]For 'From the Netherlands' and 'Birth year' less observations are available. For the grant winners, there are 566 and 517 observations respectively. For the grant losers there are 543 and 487 observations.

[22]Using a regression discontinuity design.

Table 4: Balancing test

	Rank first application round > than cutoff
From the Netherlands	0.0467
	(0.0301)
Birth year	0.711**
	(0.320)
Is female	0.138***
	(0.0449)
Publications 2 - 0 years before application	0.0340
	(0.655)
MNCS 2 - 0 years before application	0.260
	(0.165)
MNJS 2 - 0 years before application	0.200*
	(0.113)
Obs	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3.3 Sample selection check

Due to the lack of information on ranks and priority scores, the sample used in the analysis is smaller than the complete dataset. In order to see whether the results from the analysis are generalizable to the entire dataset, I test whether the estimation sample and the original dataset have significantly different observable characteristics. The results are shown in Table 5.

Table 5: Sample selection check

	Original = (1) Sample = (2)				
	obs(1)	obs(2)	mean(1)	mean(2)	p-value
Ever received the grant	4638	1121	0.25	0.51	0.00000
Normalized rank	1599	1121	-5.45	-4.29	0.02071
From the Netherlands	4506	1109	0.92	0.91	0.10802
Birth year	3963	1004	1973.25	1972.85	0.02161
Female	4638	1121	0.40	0.34	0.00127
N	5759				

In percentage terms, the estimation sample has more winners than the original data. This indicates that my analysis is looking at a group closer around the cutoff. Also, the estimation sample is a bit more divers; it has a one percentage point less people who apply for the grant from the Netherlands. Finally, the applicants in the sample are on average younger and less likely to be female.

The differences between the two datasets are significant but small for (almost) all of the listed control variables. Therefore the results from my analysis can be considered generalizable.

6 Results

6.1 T-test

When Bol et al. (2018) analyzed the effect of the Veni-grant on research productivity, they conducted a t-test using a bandwidth of -2 ranks below to +2 ranks above the cutoff. The productivity indicators they tested were publication and citation-scores as well as the H-index. The results showed a higher productivity of grant winners compared to that of grant losers. The differences were not significant at the five percent level.

In order to see whether I find similar results with my data and my productivity indicators I conduct my own t-tests with the same bandwidth. Like Bol et al. I take the natural log of the productivity indicators in order to normalize the left-skewed distributions and then normalize the productivity indicators across rounds and domains using a z-score in order to control for differences between publication- and citation cultures of scientific disciplines and generations.[23] Table 6 shows the results from the t-test with the -2/+2 bandwidth.

My findings are comparable to those of Bol et al. (2018). My results suggest that the grant winners were more productive in terms of all productivity indicators. These grant effects are statistically significant at the 10% level for the short-term TNCS and the long-term MNCS, TNCS and TNJS. However, for all productivity indicators, the results fail to provide statistically significant evidence that the grant had an effect at the 5% level. When using a regression discontinuity design in my main analysis I am able to make use of a larger amount of observations, which allows me to estimate the effect more precisely.

[23] Bol et al. (2018) did not normalize across rounds, so in that sense our t-tests differ.

Table 6: T-test productivity indicators (bw -2/+2)

	Grant loser = (1) Grant winner = (2)				
	obs(1)	obs(2)	mean(1)	mean(2)	p-value
0-4 years after application					
Normalized Publications	50	97	-0.15	0.00	0.34794
Normalized MNCS	50	97	-0.16	0.06	0.20423
Normalized MNJS	50	97	-0.05	0.01	0.72475
Normalized TNCS	50	97	-0.24	0.08	0.05431
Normalized TNJS	50	97	-0.13	0.03	0.30088
5-8 years after application					
Normalized Publications	50	97	-0.08	0.11	0.24837
Normalized MNCS	50	97	-0.15	0.14	0.09097
Normalized MNJS	50	97	-0.13	0.13	0.12403
Normalized TNCS	50	97	-0.18	0.12	0.07323
Normalized TNJS	50	97	-0.20	0.11	0.06247
N	147				

6.2 First stage

IV estimates are only reliable when there is a meaningful first stage relationship between the instrument and the independent variable of interest. I check the strength of my analysis' first stage by regressing the chance of ever receiving a grant on a dummy indicating whether a person's first Veni application rank was above or below the cutoff. I thereby include all the controls which are also used in the second stage[24]. In order to check whether the priority score really is a worse predictor of grant receival I also include the first stage regression using the score as a running variable. I dropped 53 observations from the analysis which had no information on the priority score, such that I could compare the adjusted R^2 from both regressions. Table 7 shows these regressions for a rank and score dummy respectively. A rank above the cutoff increases the chance of ever receiving a grant by 71%. A priority score below the cutoff increases the chance of ever receiving a grant by 56%.[25] These estimates as well as a higher adjusted R^2 and F-statistic when using the rank indicate that the first stage is stronger when using the rank, which is an argument in favor of using it as a running variable.

[24] First stage regressions which include MNCS or MNJS from 2 to 0 years before the application show similar estimates and adjusted R^2's.

[25] A lower priority score is 'better'.

Table 7: First Stage Rank and Score

	Received Veni grant ever (Rank)	Received Veni grant ever (Score)
Rank/Score first application round > than cutoff	0.734***	-0.561***
	(0.0325)	(0.0380)
Rank/Score normalized across competitions	0.00443***	-0.105***
	(0.00127)	(0.0157)
Applicant is female	0.0273	0.0485**
	(0.0203)	(0.0229)
Applicant is from the Netherlands	0.0334	0.0393
	(0.0320)	(0.0358)
Applicant's year of birth	0.0101***	0.0126***
	(0.00235)	(0.00280)
Publications from 2 to 0 years before the application	0.00372	0.00412
	(0.00249)	(0.00265)
CWTS tool used information on applicant's e-mail address	0.0978***	0.146***
	(0.0237)	(0.0296)
round	Yes	Yes
domain	Yes	Yes
Observations	965	965
Adjusted R^2	0.712	0.631
F	300.5	184.6

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.3 Fuzzy Regression Discontinuity

As stated before, I am interested in the grant effect on productivity over two periods, one being the short-term period following the grant acquisition, the other being the long-term period after the end of the Veni grant funding. The short-term productivity indicators are summed over four years starting from the moment the Veni grant was received or applied for. The long-term productivity indicators are summed over four years covering five to eight years after the grant application.

First of all, I look at a visual representation of the relationship between productivity and the rank to see if research productivity is discontinuous at the cutoff. Figures 3 and 4 show the short-term and long-term log-, domain- and round-normalized publications as a function of the normalized rank. With the naked eye, no discontinuity in productivity at the cutoff rank can be observed from these figures. The same is valid for graphical depictions of other normalized productivity indicators (the MNCS and the MNJS) as a function of the normalized rank. These figures (5-8) can be found in Section 9.2 of the Appendix.

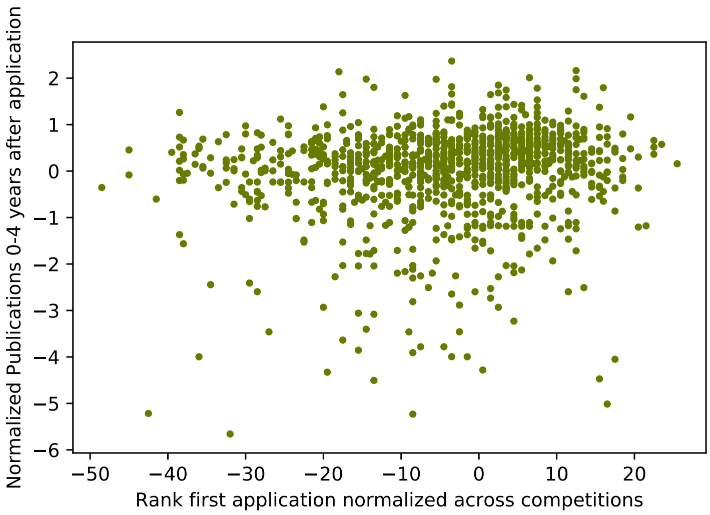

Figure 3: Normalized publications from 0-4 years after application as a function of the normalized rank

In Table 8 I show the 2sls estimation results using a short-term and long-term version of five different productivity indicators as dependent variables; publications, the MNCS, the MNJS, the TNCS and the TNJS. For each of these productivity indicators, grant estimates are shown from regression results from six model specifications that each include more control variables on top of the previous specification. In the first model specification I run the most basic 2sls estimation including the rank and the dummy indicating whether I

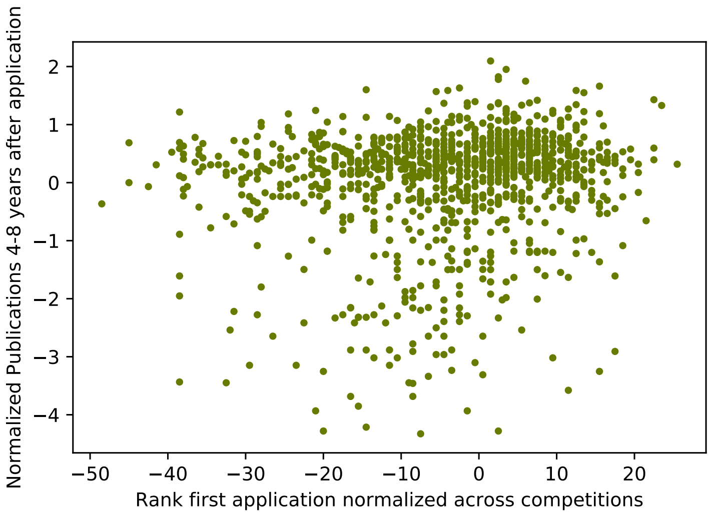

Figure 4: Normalized publications from 5-8 years after application as a function of the normalized rank

had the applicant's e-mail address for the CWTS tool. In the second I control for round and domain fixed effects. To the third specification I add the control variables which I showed in Section 5.3.2 to be discontinuous at the cutoff, namely gender and year of birth.[26] The fourth specification includes the other two control variables amongst which the productivity indicator from two years before the application. The fifth and sixth specification include a second and third degree polynomial of the running variable respectively. In Table 8, only grant estimates and their standard errors are shown. The complete regression tables including the coefficients of the control variables are shown in Section 9.3.

For all grant estimates standard errors decrease while moving from the first to the second model specification. This is not surprising since the normalization of the rank is not fully complete without controlling for the rounds and domains. The standard errors increase when going from the second to the third model specification. They decrease again with the inclusion of more controls and start increasing again when higher degree polynomials are added to the regression.

For all productivity indicators, short- and long-term, I find that the addition of the variables that are discontinuous at the cutoff, gender and birth year, increases the grant estimate. When the other two controls are included in model (4), the dummy indicating whether the applicant is from the Netherlands and the productivity indicator summed over the two years prior to the application, the grant estimate decreases again. This suggests that the prior productivity of

[26] I chose not to include the MNJS from two years before the application because it was considered discontinuous only at the ten percent significance level.

the applicants is also discontinuous at the cutoff.

The grant estimate is positive for all short- and long-term productivity indicators. Exceptions are the grant estimates for the short-term TNCS and TNJS estimated using model (6). For all productivity indicators, the grant coefficients are larger for the long-term period than the short-term period. In the short term, the Veni grant increases the number of publications by a value between $0.4 - 1.1$, depending on the assumed functional form. The short-term increase ranges between $0.13 - 0.22$ for the MNCS and between $0.10 - 0.23$ for the MNJS.[27] In the long term, the Veni grant increases the number of publications by a value ranging between $3.7 - 4.8$, the MNCS by a value ranging between $0.10 - 0.50$ and the MNJS by a value ranging between $0 - 0.35$.

The estimated effect of the grant using model (4) is statistically significant at the five percent level for the short-term MNJS and for all long-term productivity indicators.

Going to model specifications (5) and (6), we can see that for all productivity indicators, the grant estimates are dropping due to the increased flexibility of the functional form. This also has the direct consequence that the statistical significance of the estimates disappears for all model (5) and (6) specifications except for the regression with long-term publications.

In these analyses I use all the available data for the regressions. When using observations so far away from the cutoff, one runs a greater risk to obtain biased estimates. Therefore, I repeat the regressions of all the treated productivity indicators using my preferred model specification (model 4) for smaller bandwidths. In Section 9.4 of the Appendix, tables 20-24 show the long-term 2sls estimates for 100, 75, 50 and 25%[28] of the data around the cutoff.

[27] I do not discuss the grant coefficients for the TNCS and the TNJS because their interpretation is ambiguous.

[28] I defined a bandwidth of x% as x% of all the data above the cutoff plus x% of all the data below the cutoff.

Table 8: 2SLS estimated effect of the grant on productivity outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Publications, 4 years after the grant application	2.297	1.161	1.701	1.125	0.681	0.431
	(1.934)	(1.838)	(1.937)	(0.861)	(0.991)	(1.012)
Publications, 5 to 8 years after the grant application	5.599**	2.813	5.488**	4.768**	4.443*	3.726
	(2.579)	(2.485)	(2.543)	(2.037)	(2.360)	(2.507)
MNCS, 4 years after the grant application	0.386**	0.276*	0.351**	0.217	0.192	0.134
	(0.151)	(0.143)	(0.156)	(0.132)	(0.153)	(0.159)
MNCS, 5 to 8 years after the grant application	0.555***	0.381*	0.585**	0.496**	0.294	0.101
	(0.212)	(0.211)	(0.228)	(0.230)	(0.183)	(0.202)
MNJS, 4 years after the grant application	0.381***	0.265**	0.319***	0.225**	0.176	0.0997
	(0.115)	(0.106)	(0.115)	(0.0983)	(0.113)	(0.117)
MNJS, 5 to 8 years after the grant application	0.463***	0.287**	0.408**	0.345**	0.170	0.00230
	(0.156)	(0.144)	(0.162)	(0.167)	(0.141)	(0.147)
TNCS, 4 years after the grant application	7.109**	4.770	5.437	1.290	0.732	-0.180
	(3.579)	(3.346)	(3.474)	(2.366)	(2.739)	(2.892)
TNCS, 5 to 8 years after the grant application	14.07***	7.843	13.28***	9.089**	7.400	5.017
	(5.209)	(5.051)	(5.035)	(4.424)	(4.884)	(5.315)
TNJS, 4 years after the grant application	6.049*	3.703	3.992	1.254	0.483	-0.270
	(3.255)	(3.059)	(3.117)	(1.674)	(1.957)	(2.031)
TNJS, 5 to 8 years after the grant application	11.97***	5.904	10.08**	7.228**	5.177	3.209
	(4.440)	(4.203)	(4.200)	(3.537)	(3.983)	(4.270)
Observations	1121	1121	1004	992	992	992

List of control variables added per regression.

(1): rank and e-mail; (2): round and domain; (3): gender and year of birth;

(4): from the Netherlands and productivity indicator 2 years before;

(5): second degree polynomial; (6): third degree polynomial.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.3.1 RD without academic drop-outs

The fuzzy regression discontinuity analysis provides evidence that the Veni grant increases the qualitative and quantitative research productivity of researchers over the long term. In this section I analyze the importance of one possible explanation for this effect. This explanation proposes that grant losers remove themselves from academia much more often than grant winners do, leading them to publish less on average. From the applicants in my estimation sample, 35 people did not publish anything during the eight years after their application.[29] Seven were grant losers, eight were grant winners. This gives an indication that most researchers continued doing research, also after getting rejected for the Veni.

When looking at publication scores in the short- and the long term separately, I find that in the short term, 36 grant losers as opposed to 30 grant winners did not publish anything. In the long term, The amount of grant losers who stop publishing increases up to 77, while the amount of grant winners who no longer publish is just 31 people. There are 31 applicants (15 grant losers, 16 grant winners) who did not publish anything in the short term, but did publish in the long term and there are 73 applicants (56 grant losers, 17 grant winners) who stopped publishing after the short term period.

These facts clarify that, after applying for a Veni, more grant losers than grant winners removed themselves from academia. This is valid for the short term, yet much more for the long term. Therefore, it is safe to infer that this mechanism at least partially causes my finding that the Veni grant positively affects research productivity. I therefore investigated the importance of this mechanism by repeating the fuzzy regression discontinuity analyses without these drop-outs. Table 9 displays the results.

When repeating the analyses without the drop-outs, the results do not change much for any of the productivity indicators in the short term. This is not surprising since only six more grant losers than grant winners have zero publications in the short-term period following their application.

In the long term, all coefficients remain of the same sign but decrease in size. The effect of the grant on publications is marginally affected. The effect of the grant on the MNCS and the MNJS, however, decreases by about 0.10 when excluding the drop-outs. With the inclusion of the drop-outs, the articles of grant winners are cited above the average of the field and year 49.6 percentage points more than the articles of grant losers. The publications from journals that these articles are published in are cited above the average of the field and year 34.5 percentage points more than the publications from journals that the grant losers' articles are published in. These coefficients change to 39.6 and 23.8 when the academic drop-outs are excluded from the analysis. This means that the articles of grant winners and the publications from their journals are cited above average about 10 percentage points more due to the fact that there are more grant losers with an MNCS and MNJS of zero than grant winners.

[29] This being according to my data. It could be that the CWTS tool just did not retrieve their publications.

Due to the drop in observations[30] the statistical significance of the grant effects disappears for the MNCS and the MNJS. Still, the coefficients suggest that the grant leads to a higher qualitative productivity, even when academic drop-outs are excluded from the analysis.

[30]In Table 9, for each model specification the first 'total observations' concerns the short-term analysis, the second 'total observations' concerns the long-term analysis.

Table 9: 2SLS estimated effect of the grant on productivity outcomes without academic drop-outs

	(1)	(2)	(3)	(4)	(5)	(6)
Publications, 4 years after the grant application	2.196	1.196	1.705	1.190	0.663	0.438
	(2.011)	(1.899)	(1.993)	(0.894)	(1.029)	(1.050)
Publications, 5 to 8 years after the grant application	4.314	1.867	4.697*	4.509**	4.350*	3.797
	(2.813)	(2.677)	(2.737)	(2.171)	(2.514)	(2.682)
MNCS, 4 years after the grant application	0.367**	0.272*	0.348**	0.217	0.179	0.120
	(0.152)	(0.145)	(0.159)	(0.136)	(0.158)	(0.164)
MNCS, 5 to 8 years after the grant application	0.420*	0.259	0.460*	0.396	0.222	0.0381
	(0.227)	(0.224)	(0.243)	(0.245)	(0.195)	(0.215)
MNJS, 4 years after the grant application	0.367***	0.261**	0.313***	0.228**	0.166	0.0861
	(0.114)	(0.107)	(0.116)	(0.101)	(0.116)	(0.121)
MNJS, 5 to 8 years after the grant application	0.334**	0.172	0.283*	0.238	0.0837	-0.0760
	(0.163)	(0.150)	(0.170)	(0.176)	(0.148)	(0.156)
TNCS, 4 years after the grant application	7.119*	5.026	5.638	1.389	0.649	-0.226
	(3.726)	(3.491)	(3.614)	(2.468)	(2.855)	(3.017)
TNCS, 5 to 8 years after the grant application	12.20**	6.438	12.25**	8.557*	7.129	4.911
	(5.673)	(5.455)	(5.423)	(4.775)	(5.253)	(5.742)
TNJS, 4 years after the grant application	6.044*	3.876	4.089	1.333	0.412	-0.320
	(3.395)	(3.179)	(3.219)	(1.750)	(2.043)	(2.124)
TNJS, 5 to 8 years after the grant application	10.17**	4.580	9.017**	6.614*	4.737	2.923
	(4.834)	(4.545)	(4.533)	(3.808)	(4.260)	(4.588)
Observations	1055/1013	1055/1013	941/903	932/895	932/895	932/895

List of control variables added per regression.
(1): rank and e-mail; (2): round and domain; (3): gender and year of birth;
(4): from the Netherlands and productivity indicator 2 years before;
(5): second degree polynomial; (6): third degree polynomial.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Conclusions and discussion

Contrary to the findings of other authors who estimate no effect of a competitive grant on research productivity (Bol et al., 2018; Lanser & van Dalen, 2013), I can infer from my analysis that there is a positive impact of receiving a Veni grant on post-Veni research productivity. In the fifth to eighth year following their application, researchers who received the grant published more, were cited more often than their peers and published in higher impact journals.

By summing the publications and their citations over two distinct periods, I cannot perfectly distinguish between short and long-term productivity, mainly because the lag with which scientific work is published can be substantial. Articles that were published in the long-term may have been written in the short term. Hence, the long-term productivity may in fact reflect short-term productivity to some extent. Even publications that I ascribe to the Veni funding may have actually been produced during the scientist's PhD. Similarly, articles that were published *after* the five to eight year period may have been produced in that period. These plausible inaccuracies may cause an over- or underestimation of the average productivity (short- and long-term) of grant winners and losers alike. However, it does not invalidate my identification strategy since there is no reason to expect this over-/underestimation to differ between losers and winners.

As was pointed out by Jacob & Lefgren (2011), the regression discontinuity design estimates the effect of the grant funding on the marginal applicant. Since the marginal applicant can be expected to shift to another source of funding after having his Veni proposal rejected, the estimated effect of the Veni grant most likely reflects the difference between the effect of the Veni and the effect of alternative funding on research productivity. In future studies, more information on the alternative funding received by grant losers should be collected in order to find out to what baseline the competitive grant funding ought to be compared.

Due to a lack of information, I did not control for whether the applicants applied for and/or received other (Dutch or foreign) grants while applying for the Veni. Therefore, the positive impact of the Veni grant could be underestimated if we were to interpret it more generally as the impact of a competitive grant.

The estimated positive effect of the grant on research productivity could be driven by a multitude of factors. In my analysis I checked how much of the effect was driven by grant losers dropping out of academia more often. My results suggest that this determinant explains the grant effect at least partially. I propose the following explanations for this higher drop-out rate and any remaining grant effects on productivity.

Missing out on the monetary advantage of the grant in the short term may have caused grant losers to drop out of academia in the long term. According to this reasoning, the lack of resources limits the grant losers' ability to perform their preferred research which in turn reduces their motivation. On top of that, the winners' productivity may have been increased due to the monetary advantage regardless of losers dropping out. Especially in the natural sciences, grant winners may be able to do better research because the grant enables them to acquire expensive machinery. As mentioned previously, the publications which I consider produced in the long term might have been written in the short-term period. Therefore, it is possible that with the use of Veni grant money some increased 'long-term' productivity of grant winners came about.

The grant winner's exemption (or even, exclusion) from teaching provides another explanation for the higher academic drop-out among grant losers. Supposedly, the grant winner would have more time to invest into research and into the development of his research skills. Subsequently, the research skills of the grant losers would in comparison not develop sufficiently over time, therefore reducing their chances of obtaining follow-up research positions.

Through a different channel, the teaching requirement would cause grant losers to drop out of academia more often because they enjoy their academic positions less since they are allowed to a lesser extent to focus on the research that they personally want to pursue. In this reasoning one can also conjecture that teaching is considered less enjoyable than research. Hence, a smaller time endowment could affect both the quality of the research(er) as well as his motivation to do research.

A perhaps more plausible explanation for the higher amount of academic drop-outs among grant losers is the signaling power of a Veni grant. The Veni is a positive signal to the research community, which leads to more exposure and opportunities, which makes the chances of acquiring a research position or tenure more likely. Researchers who receive the grant have been shown to benefit from the Mattheus effect (Bol et al. (2018)), which means that they are more likely to receive a follow-up grant, usually the Vidi or the Vici, after receiving the Veni. Therefore, the estimated long-term effect of receiving a Veni grant in part extends itself to the effects that future grants have on research productivity. This may explain both the lower drop-out rate among grant winners and their higher long-term productivity.

The acquisition of knowledge and skills can be considered a justified cause for the grant, while signaling does not add to the productive potential of the researcher. My findings corroborate the idea that it is rather the demotivation or inability of rejected grant applicants to *continue* doing research which drives the positive effect of the Veni grant on research productivity. This implies that the grant does in fact not genuinely enhance the productivity of researchers. Still, no hard conclusions can be made about the estimated effect's underlying mechanism, therefore it is important to disentangle between all of the proposed explanations in further research.

8 Bibliography

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. Economic Policy, 25(61), 7-59.

Auranen, O., & Nieminen, M. (2010). University research funding and publication performanceAn international comparison. Research Policy, 39(6), 822-834.

Bloch, C., Graversen, E. K., & Pedersen, H. S. (2014). Competitive research grants and their impact on career performance. Minerva, 52(1), 77-96.

Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. Proceedings of the National Academy of Sciences, 115(19), 4887-4890.

Bornmann, L., Wallon, G., & Ledin, A. (2008). Does the committee peer review select the best applicants for funding? An investigation of the selection process for two European molecular biology organization programmes. PLoS ONE, 3(10), e3480.

Bornmann, L., Leydesdorff, L., & van den Besselaar, P. (2010). A Meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. Journal of Informetrics, 4(3), 211220.

Campbell, D., Picard-Aitken, M., Cote, G., Caruso, J., Valentim, R., Edmonds, S., et al. (2010). Bibliometrics as a performance measurement tool for research evaluation: The case of research funded by the National Cancer Institute of Canada. American Journal of Evaluation, 31, 6683.

Caron, E., & Van Eck, N.J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. In Proceedings of the 19th International Conference on Science and Technology Indicators (pp. 79-86).

CWTS Leiden Ranking. (n.d.). Fields. Retrieved from
http://www.leidenranking.com/information/fields#micro-level-fields

Danthi, N. S., Wu, C. O., Shi, P., & Lauer, M. (2014). Percentile ranking and citation impact of a large cohort of NIHLBI-funded cardiovascular R01grants.Circulation Research, 114, 600606.

Decullier, E., Huot, L., & Chapuis, F. R. (2014). Fate of protocols submitted to a French national funding scheme: A cohort study. PLoS ONE, 9(6), e99561.http://dx.doi.org/10.1371/journal.pone.0099561

Gallo, S. A., Carpenter, A. S., Irwin, D., McPartland, C. D., Travis, J., Reynders, S., et al. (2014). The validation of peer review through research impact measures and the implications for funding strategies. PLoS ONE, 9(9), e106474. http://dx.doi.org/10.1371/journal.pone.0106474

Gerritsen, S., Plug, E., & van der Wiel, K. (2013). Up or out? How individual research grants affect academic careers in the Netherlands (No. 249). CPB Netherlands Bureau for Economic Policy Analysis.

Hornbostel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & von Ins, M. (2009). Funding of young scientist and scientific excellence. Scientometrics, 79, 171190.

Huffman, W. E., & Evenson, R. E. (2006). Do formula or competitive grant funds have greater impacts on state agricultural productivity?. American Journal of Agricultural Economics, 88(4), 783-798.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. Journal of public economics, 95(9-10), 1168-1177.

Kaltman, J. R., Evans, F. J., Danthi, N. S., Wu, C. O., DiMichele, D. M., & Lauer, M. S. (2014). Prior publication productivity, grant percentile ranking andtopic-normalized citation impact of NIHLBI-funded cardiovascular R01grants. Circulation Research, 115, 617624.

Lanser, D., & van Dalen, R. (2013). The effects of research grants on scientific productivity and utilisation. CPB Netherlands Bureau for Economic Policy Analysis.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of economic literature, 48(2), 281-355.

Mutz, R., Bornmann, L., & Daniel, H.-D. (2014). Testing for fairness and predictive validity of research funding decisions: A multi-level multiple imputation for missing data approach using ex-ante and ex-post peer evaluation data from the Austrian Science Fund. Journal of the Association for Information Science and Technology, http://dx.doi.org/10.1002/asi.23315 (published online)

Nyström, P., Noijons, E., & Nane, T. (2014). Bibliometric analysis of Mälardalen University. Centre for Science and Technology Studies, Leiden University.

Reinhart, M. (2009). Peer review of grant applications in biology and medicine. Reliability, fairness, and validity. Scientometrics, 81(3), 789809.

Saygitov, R. T. (2014). The impact of funding through the FR presidents grants for young scientists on research productivity: A quasi-experimental studyand a brief systematic review. PLoS ONE, 9(1), e86969. http://dx.doi.org-/10.1371/journal.pone.0086969

Staiger, D. O., & Stock, J. H. (1994). Instrumental variables regression with weak instruments.

Stephan, P. E. (2010). The economics of science. In Handbook of the Economics of Innovation (Vol. 1, pp. 217-273). North-Holland.

van Dalen, R., van Elk, R., & van Vuuren, D. (2015). Publieke onderzoeksfinanciering: de voor- en nadelen van verschillende financieringsmethoden. CPB Netherlands Bureau for Economic Policy Analysis.

van den Besselaar, P. A. A., Hinze, S., & Lottmann, A. (2013). More competition, better science? The predictive validity of grant decisions. STI proceedings, (18).

van den Besselaar, P., & Sandström, U. (2015). Early career grants, performance, and careers: A study on predictive validity of grant decisions. Journal of Informetrics, 9(4), 826-838.

van Leeuwen, T. N., & Moed, H. F. (2012). Funding decisions, peer review, and scientific excellence in physical sciences, chemistry, and geosciences. Research Evaluation, 21, 189198.

9 Appendix

9.1 Data collection procedure

In order for the CWTS tool to work effectively I needed to provide it with information on the applicants personal details. This data was readily available at the CPB in dta-format for the years 2002 to 2008. These variables included:

1. Email address
2. The associated research institute
3. Subdivision

I created a new file specifically designed for the author disambiguation tool. I dropped all observations for which there were no priority scores, since at the time I thought that I was going to do my analysis using the priority score as the running variable. Later I found that this variable was not perfect as a running variable so I decided to also use the rank.[31] I then removed second applications from authors in order for the tool not to look for the same author twice. This I did by identifying individuals by grouping their last name, initials, surname prefix(es) and date of birth and removing duplicates from this identification variable. Finally, I dropped most variables from the file and sent the CWTS the remainder containing the following information:

1. Last name
2. Initials
3. Surname prefix(es)
4. Email address — 187/1,740 missing
5. The associated research institute — 14/1,740 missing
6. Subdivision — 351/1,740 missing

For 1,553 out of 1,740 applicants, information was available for the first 5 variables from the list. For 1,276 out of 1,740 applicants information was available on subdivision as well.

The data was then fed to the tool, which uses an author disambiguation method for large bibliographic databases that uses rule-based scoring and clustering. Its method relies on five steps:

1. Author names (last name, surname prefix(es) and initials) are grouped into blocks.
2. Publications are paired within these blocks based on the assignment of points (which from here on I will refer to as the matching score) from four categories of scoring rules that are based on bibliographic metadata. These categories include author-, article-, source- and citation rules. The publications matching scores need to surpass a predefined threshold in order for them to be paired. The height of this threshold is dependent on the number of publications present within the block. The larger the block size, the higher the threshold needs to be since the homonym problem is logically more problematic.
3. The publication pairs will be clustered through single-linkage clustering if their matching scores surpass the predefined threshold. In this single-linkage clustering, all publication pairs start out as their own cluster and are then sequentially combined with the cluster that has the highest matching score up until the point where the matching scores between the clusters no longer are high enough for them to be combined. The usage of the scoring rules in the second and third stage is meant to optimally address the homonym problem.

[31]From the 4,013 proposals that had no priority score, there were 34 that did have a rank but were nonetheless (and unfortunately) dropped from my dataset.

4. Some of these clusters are linked across blocks based on matching e-mail addresses. This fourth step aims to solve the synonym problem as well as possible.

5. All the resulting clusters are assigned a matching score with the authors from my dataset based on scoring rules for all the variables in my datafile (see the list). The clusters are then included in my verification list.

A precise explanation of the method and its underlying rules described above can be found in Caron and van Eck (2014). These authors also evaluated their tool using two datasets with mainly Dutch researchers.

After the data went through the process described above, I looked at the resulting clusters myself and verified them manually in excel by looking at the authors initials, last names, e-mail addresses, their related organizations and the scientific subject(s) to which their proposal was linked. For a given author from my dataset I checked per cluster for each of these variables whether the values from the dataset and the cluster matched. According to a set of rules I decided whether the cluster belonged to the author. Section 9.1.1 describes these rules in detail.

9.1.1 Cluster verification method

Once the author verification tool from the CWTS had created clusters of publications for all authors in my dataset, I manually checked for each cluster from each author from my dataset whether the variable values from the dataset matched the variable values from the cluster. According to the following set of rules I decided whether the cluster belongs to the author.

- If the last name did not match, the cluster was rejected as belonging to the author.
- If the last name did match then the following happened:
  - If the cluster contained different fully written out names for a given initial that were clearly not a result of misspelling, I concluded that it contained publications for more than one author and therefore I rejected the cluster.
  - If this was not the case, the following happened:
    - If the e-mail address matched[32] [33], the following happened:
      - If the initials did not mismatch, the cluster was accepted.[34]

[32]If the content of the e-mail address matched, but the domain of the e-mail address in the cluster was written differently, the e-mail addresses were considered to match when the domains were one anothers synonyms.

[33]If the dataset contained no e-mail address for a researcher profile, the e-mail address from the cluster was still considered to match when it matched an e-mail address found in a cluster which had been previously accepted based on a matching last name and multiple initials.

[34]If the email address in my dataset contained a different first name or initial than the first name/initial, clusters containing that initial/first name were also considered to match the first name of the author.

      - If the initials mismatched, I checked on google and LinkedIn if the researcher is also known under the other first name. If that was the case, the cluster was accepted.
      - Otherwise, the cluster was rejected.
   * If the e-mail address did not match, the following happened:
      - If there was only one initial available in the dataset, the following happened:
         - If the organization and scientific subject(s) matched, the cluster was accepted.
         - If the scientific subject(s) matched but the organization(s) did not match, I checked on LinkedIn and google if the researcher had filled a position at the institutions from the cluster. If he had, the cluster was accepted.
         - Otherwise, the cluster was rejected.
      - If the first letter of the first name or the first initial in the cluster mismatched the first initial from my data and/or the first letter of the second name or the second initial in the cluster mismatched the second initial from my data, the following happened:
         - If the organization and field did match and there was another cluster with matching initials/first name, last name, organization and field, I checked on google and LinkedIn if the researcher is also known under the other first name. If that was the case, the cluster was accepted.
         - Otherwise, the cluster was rejected.
      - If there were two or more initials in the dataset and the first letter of the first name or the first initial in the cluster did match the first initial from my data, the following happened:
         - If at least the first two initials matched, while mismatches for the other initials were only due to a lack of initials in the cluster, the cluster was accepted.
         - If there was no second name or initial available in the cluster, the following happened:
            · If the second initial of the dataset was found in the first letter of the second name or the second initial in the e-mail address provided by the cluster, the cluster was accepted.
            · If the organization and scientific subject(s) matched and the other available clusters for this author were clearly rejected, the cluster was accepted.
            · If the organization did not match and there was another cluster with matching initials/first name, last name and whose organization did match the dataset, I checked on

LinkedIn and google if the researcher had filled a position at both institutions. If he had, the cluster was accepted.

– All other clusters were rejected.

9.2 Normalized productivity indicators as a function of the normalized rank

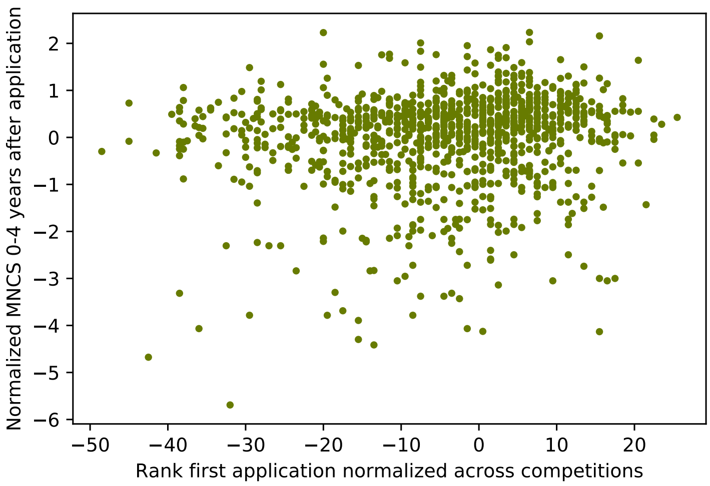

Figure 5: Normalized MNCS from 0-4 years after application as a function of the normalized rank

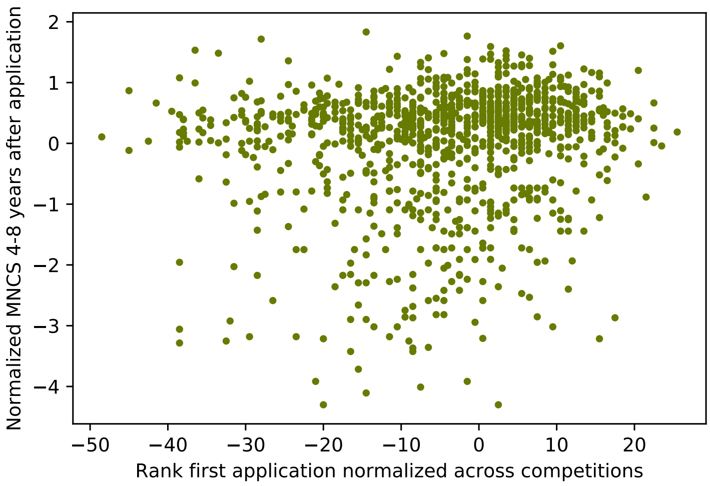

Figure 6: Normalized MNCS from 5-8 years after application as a function of the normalized rank

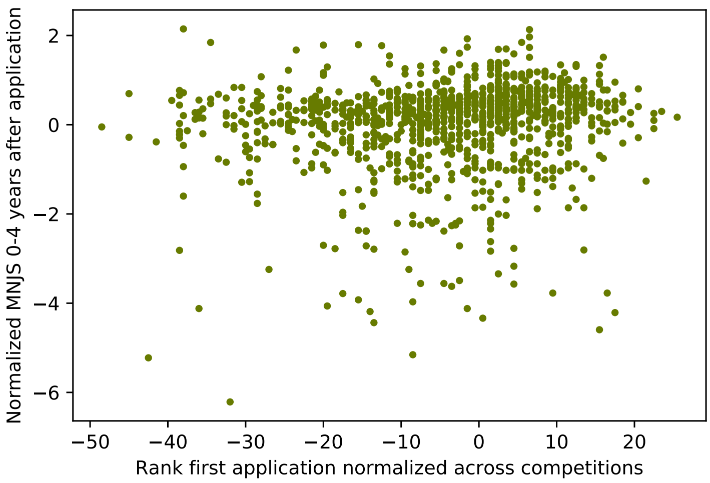

Figure 7: Normalized MNJS from 0-4 years after application as a function of the normalized rank

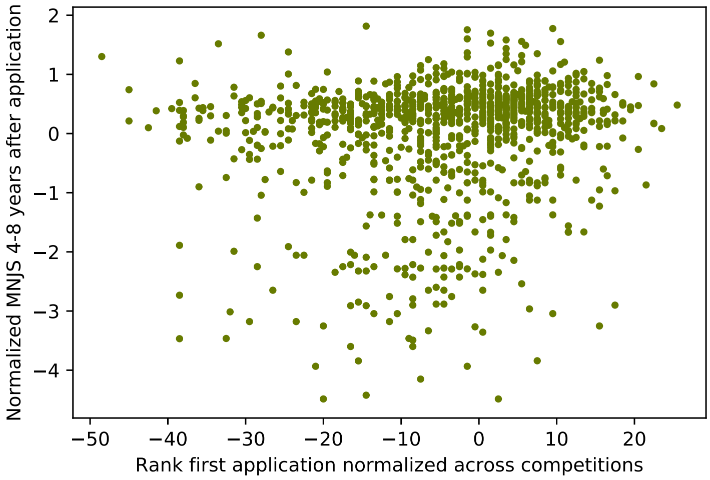

Figure 8: Normalized MNJS from 5-8 years after application as a function of the normalized rank

9.3 2SLS estimation with the rank as the running variable

Here I show the full 2SLS regression tables with six model specifications from which the grant estimates were displayed in Table 8. In each table, the title refers to the outcome variable used.

Table 10: Publications over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	2.297	1.161	1.701	1.125	0.681	0.431
	(1.934)	(1.838)	(1.937)	(0.861)	(0.991)	(1.012)
Rank normalized across competitions	-0.00988	0.0568	0.0954	-0.0207	0.0107	0.0280
	(0.0627)	(0.0626)	(0.0614)	(0.0299)	(0.0475)	(0.0489)
CWTS tool used information on applicant's e-mail address	3.856***	4.042***	3.379***	1.460**	1.491**	1.530**
	(0.981)	(1.011)	(0.998)	(0.637)	(0.638)	(0.641)
Applicant is female			-2.847***	-0.840*	-0.827*	-0.814*
			(0.539)	(0.439)	(0.439)	(0.439)
Applicant's year of birth			0.0818	-0.0238	-0.0208	-0.0196
			(0.0809)	(0.0551)	(0.0544)	(0.0546)
Applicant is from the Netherlands				0.418	0.431	0.420
				(0.595)	(0.593)	(0.592)
Publications from 2 to 0 years before the application				1.768***	1.770***	1.770***
				(0.130)	(0.129)	(0.129)
(Normalized rank)^2					0.00120	-0.0000851
					(0.00130)	(0.00197)
(Normalized rank)^3						-0.0000502
						(0.0000493)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: MNCS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.386**	0.276*	0.351**	0.217	0.192	0.134
	(0.151)	(0.143)	(0.156)	(0.132)	(0.153)	(0.159)
Rank normalized across competitions	-0.00450	0.00605	0.00586	-0.00109	0.000681	0.00467
	(0.00432)	(0.00430)	(0.00509)	(0.00419)	(0.00636)	(0.00706)
CWTS tool used information on applicant's e-mail address	-0.0576	-0.0238	-0.112	-0.0826	-0.0806	-0.0714
	(0.194)	(0.186)	(0.195)	(0.166)	(0.167)	(0.168)
Applicant is female			-0.218***	-0.147**	-0.146**	-0.143**
			(0.0822)	(0.0696)	(0.0696)	(0.0695)
Applicant's year of birth			0.00854	0.00248	0.00265	0.00292
			(0.0112)	(0.0107)	(0.0107)	(0.0108)
Applicant is from the Netherlands				0.0409	0.0415	0.0390
				(0.0946)	(0.0945)	(0.0945)
MNCS from 2 to 0 years before the application				0.380***	0.380***	0.381***
				(0.0393)	(0.0391)	(0.0392)
(Normalized rank)ˆ2					0.0000680	-0.000229
					(0.000162)	(0.000251)
(Normalized rank)ˆ3						-0.0000116
						(0.00000753)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: MNJS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.381***	0.265**	0.319***	0.225**	0.176	0.0997
	(0.115)	(0.106)	(0.115)	(0.0983)	(0.113)	(0.117)
Rank normalized across competitions	-0.00431	0.00481	0.00531	-0.000835	0.00253	0.00781
	(0.00355)	(0.00349)	(0.00414)	(0.00364)	(0.00505)	(0.00560)
CWTS tool used information on applicant's e-mail address	0.0395	0.0484	-0.0187	-0.0106	-0.00689	0.00530
	(0.107)	(0.106)	(0.108)	(0.0944)	(0.0941)	(0.0945)
Applicant is female			-0.216***	-0.147***	-0.146***	-0.142***
			(0.0596)	(0.0526)	(0.0524)	(0.0521)
Applicant's year of birth			0.0162**	0.00767	0.00796	0.00832
			(0.00707)	(0.00651)	(0.00648)	(0.00658)
Applicant is from the Netherlands				-0.0501	-0.0487	-0.0520
				(0.0873)	(0.0872)	(0.0869)
MNJS from 2 to 0 years before the application				0.358***	0.360***	0.361***
				(0.0308)	(0.0306)	(0.0307)
(Normalized rank)ˆ2					0.000130	-0.000263
					(0.000151)	(0.000205)
(Normalized rank)ˆ3						-0.0000153**
						(0.00000716)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: TNCS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	7.109**	4.770	5.437	1.290	0.732	-0.180
	(3.579)	(3.346)	(3.474)	(2.366)	(2.739)	(2.892)
Rank normalized across competitions	-0.104	0.0482	0.159	-0.00539	0.0338	0.0967
	(0.117)	(0.116)	(0.107)	(0.0737)	(0.112)	(0.123)
CWTS tool used information on applicant's e-mail address	4.439**	5.673***	4.367**	2.924*	2.963*	3.106*
	(2.050)	(1.941)	(1.967)	(1.694)	(1.701)	(1.721)
Applicant is female			-5.433***	-2.701***	-2.684***	-2.638**
			(1.198)	(1.044)	(1.042)	(1.045)
Applicant's year of birth			0.160	0.0915	0.0953	0.0998
			(0.150)	(0.117)	(0.116)	(0.117)
Applicant is from the Netherlands				0.864	0.878	0.838
				(1.707)	(1.700)	(1.702)
TNCS from 2 to 0 years before the application				1.508***	1.510***	1.511***
				(0.191)	(0.191)	(0.190)
(Normalized rank)ˆ2					0.00150	-0.00318
					(0.00265)	(0.00405)
(Normalized rank)ˆ3						-0.000182
						(0.000121)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: TNJS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	6.049*	3.703	3.992	1.254	0.483	-0.270
	(3.255)	(3.059)	(3.117)	(1.674)	(1.957)	(2.031)
Rank normalized across competitions	-0.0850	0.0550	0.163*	-0.0262	0.0280	0.0802
	(0.114)	(0.113)	(0.0979)	(0.0546)	(0.0830)	(0.0905)
CWTS tool used information on applicant's e-mail address	4.575***	5.276***	4.167***	2.380**	2.433**	2.552**
	(1.466)	(1.593)	(1.587)	(1.104)	(1.104)	(1.114)
Applicant is female			-4.782***	-1.849**	-1.825**	-1.788**
			(0.951)	(0.800)	(0.797)	(0.797)
Applicant's year of birth			0.201*	0.0426	0.0477	0.0514
			(0.119)	(0.0885)	(0.0876)	(0.0882)
Applicant is from the Netherlands				0.596	0.619	0.586
				(1.170)	(1.165)	(1.164)
TNJS from 2 to 0 years before the application				1.596***	1.599***	1.599***
				(0.152)	(0.151)	(0.151)
(Normalized rank)^2					0.00209	-0.00179
					(0.00219)	(0.00319)
(Normalized rank)^3						-0.000151
						(0.0000967)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Publications over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	5.599**	2.813	5.488**	4.768**	4.443*	3.726
	(2.579)	(2.485)	(2.543)	(2.037)	(2.360)	(2.507)
Rank normalized across competitions	-0.0536	0.0653	0.0111	-0.0972	-0.0742	-0.0245
	(0.0816)	(0.0824)	(0.0811)	(0.0664)	(0.0995)	(0.111)
CWTS tool used information on applicant's e-mail address	6.227***	6.650***	5.026***	3.340**	3.362**	3.476**
	(1.608)	(1.616)	(1.622)	(1.589)	(1.591)	(1.609)
Applicant is female			-3.655***	-1.806**	-1.796**	-1.761**
			(0.926)	(0.866)	(0.870)	(0.872)
Applicant's year of birth			0.118	0.0341	0.0362	0.0398
			(0.130)	(0.115)	(0.115)	(0.116)
Applicant is from the Netherlands				2.777**	2.786**	2.754**
				(1.295)	(1.295)	(1.291)
Publications from 2 to 0 years before the application				1.635***	1.637***	1.637***
				(0.113)	(0.114)	(0.113)
(Normalized rank)ˆ2					0.000884	-0.00281
					(0.00245)	(0.00354)
(Normalized rank)ˆ3						-0.000144
						(0.000106)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: MNCS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.555***	0.381*	0.585**	0.496**	0.294	0.101
	(0.212)	(0.211)	(0.228)	(0.230)	(0.183)	(0.202)
Rank normalized across competitions	-0.0128	-0.000755	-0.00621	-0.00903	0.00509	0.0185*
	(0.00814)	(0.00806)	(0.0102)	(0.00988)	(0.00747)	(0.0101)
CWTS tool used information on applicant's e-mail address	0.367**	0.363***	0.255*	0.292**	0.307**	0.338**
	(0.151)	(0.140)	(0.143)	(0.144)	(0.143)	(0.145)
Applicant is female			-0.205**	-0.171*	-0.165*	-0.156*
			(0.0932)	(0.0921)	(0.0907)	(0.0900)
Applicant's year of birth			0.0176	0.0166	0.0179	0.0188*
			(0.0110)	(0.0112)	(0.0110)	(0.0110)
Applicant is from the Netherlands				0.102	0.107	0.0989
				(0.213)	(0.211)	(0.210)
MNCS from 2 to 0 years before the application				0.181***	0.186***	0.187***
				(0.0404)	(0.0405)	(0.0401)
(Normalized rank)^2					0.000543	-0.000451
					(0.000394)	(0.000358)
(Normalized rank)^3						-0.0000387**
						(0.0000161)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: MNJS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.463***	0.287**	0.408**	0.345**	0.170	0.00230
	(0.156)	(0.144)	(0.162)	(0.167)	(0.141)	(0.147)
Rank normalized across competitions	-0.00950	0.00184	-0.00101	-0.00360	0.00866	0.0202***
	(0.00615)	(0.00571)	(0.00733)	(0.00705)	(0.00578)	(0.00675)
CWTS tool used information on applicant's e-mail address	0.333***	0.329***	0.252**	0.276***	0.289***	0.316***
	(0.110)	(0.103)	(0.104)	(0.104)	(0.103)	(0.105)
Applicant is female			-0.187***	-0.152**	-0.146**	-0.138*
			(0.0718)	(0.0721)	(0.0710)	(0.0707)
Applicant's year of birth			0.0149*	0.0115	0.0126	0.0133
			(0.00860)	(0.00881)	(0.00868)	(0.00880)
Applicant is from the Netherlands				-0.0268	-0.0216	-0.0288
				(0.187)	(0.184)	(0.184)
MNJS from 2 to 0 years before the application				0.178***	0.185***	0.187***
				(0.0315)	(0.0312)	(0.0305)
(Normalized rank)^2					0.000473*	-0.000385
					(0.000270)	(0.000267)
(Normalized rank)^3						-0.0000334***
						(0.00000848)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: TNCS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	14.07***	7.843	13.28***	9.089**	7.400	5.017
	(5.209)	(5.051)	(5.035)	(4.424)	(4.884)	(5.315)
Rank normalized across competitions	-0.311*	-0.0228	-0.0819	-0.223	-0.105	0.0599
	(0.177)	(0.177)	(0.179)	(0.169)	(0.222)	(0.261)
CWTS tool used information on applicant's e-mail address	7.110	8.187*	4.002	2.591	2.709	3.083
	(4.381)	(4.282)	(4.368)	(4.660)	(4.683)	(4.779)
Applicant is female			-7.337***	-4.939**	-4.887**	-4.767**
			(2.095)	(2.061)	(2.063)	(2.076)
Applicant's year of birth			0.383	0.355	0.367	0.378
			(0.250)	(0.238)	(0.239)	(0.239)
Applicant is from the Netherlands				5.591*	5.634*	5.528*
				(3.099)	(3.088)	(3.071)
TNCS from 2 to 0 years before the application				1.323***	1.329***	1.331***
				(0.244)	(0.244)	(0.244)
(Normalized rank)^2					0.00456	-0.00768
					(0.00642)	(0.00783)
(Normalized rank)^3						-0.000477
						(0.000303)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: TNJS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	11.97***	5.904	10.08**	7.228**	5.177	3.209
	(4.440)	(4.203)	(4.200)	(3.537)	(3.983)	(4.270)
Rank normalized across competitions	-0.261	0.0174	-0.0289	-0.195	-0.0508	0.0854
	(0.159)	(0.155)	(0.149)	(0.132)	(0.185)	(0.209)
CWTS tool used information on applicant's e-mail address	6.811*	7.548**	4.070	2.389	2.530	2.841
	(3.829)	(3.828)	(3.921)	(4.117)	(4.141)	(4.222)
Applicant is female			-6.239***	-3.628**	-3.564**	-3.467**
			(1.821)	(1.753)	(1.748)	(1.761)
Applicant's year of birth			0.349*	0.233	0.246	0.256
			(0.205)	(0.190)	(0.191)	(0.192)
Applicant is from the Netherlands				4.784*	4.846*	4.758*
				(2.602)	(2.588)	(2.577)
TNJS from 2 to 0 years before the application				1.413***	1.420***	1.421***
				(0.137)	(0.137)	(0.137)
(Normalized rank)ˆ2					0.00555	-0.00458
					(0.00504)	(0.00650)
(Normalized rank)ˆ3						-0.000395*
						(0.000208)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1121	1121	1004	992	992	992

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

9.4 Smaller bandwidth 2SLS regressions with the rank as the running variable

Moving from 100 to 75% of the data, all grant coefficients decrease substantially for all productivity indicators. For the MNCS and MNJS, the signs even become negative. When moving from 75 to 50% of the data, the coefficient increases again for publications, but decreases further for the MNCS and MNJS. At 25% of the data, the coefficients increase again for all productivity indicators, yet remain lower than those found using 100% of the data and are even very close to zero for the MNCS and MNJS. These results suggest that the effect of the grant on qualitative research productivity may not be positive (or even negative) after all and that the results from the main analysis are biased. However, the results could also be interpreted as meaningless since the estimates from these smaller bandwidth regressions are not statistically significant for any of the productivity indicators due to the loss of observations. Therefore they cannot provide us with hard evidence on the biasedness of the full sample regressions. Figures 3-8 visually support the notion that the observations of productivity are too dispersed to get reliable estimates with so little data.

Table 20: Publications smaller bandwidth regressions

	100%	75%	50%	25%
Received Veni grant ever	4.768**	2.687	3.122	3.147
	(2.037)	(2.557)	(4.077)	(4.249)
Rank normalized across competitions	-0.0972	0.0473	0.0315	0.0401
	(0.0664)	(0.115)	(0.302)	(0.279)
Applicant is female	-1.806**	-2.011**	-3.181**	-3.712**
	(0.866)	(0.978)	(1.341)	(1.641)
Applicant is from the Netherlands	2.777**	2.649*	1.471	2.740
	(1.295)	(1.498)	(1.923)	(2.106)
Applicant's year of birth	0.0341	0.0664	0.0202	0.00165
	(0.115)	(0.121)	(0.166)	(0.175)
Publications from 2 to 0 years before the application	1.635***	1.652***	1.691***	1.673***
	(0.113)	(0.114)	(0.109)	(0.115)
CWTS tool used information on applicant's e-mail address	3.340**	4.765***	4.732***	5.158**
	(1.589)	(1.303)	(1.796)	(2.153)
round	Yes	Yes	Yes	Yes
domain	Yes	Yes	Yes	Yes
Observations	992	803	539	427

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: MNCS smaller bandwidth regressions

	100%	75%	50%	25%
Received Veni grant ever	0.496**	-0.0387	-0.140	-0.0205
	(0.230)	(0.199)	(0.305)	(0.316)
Rank normalized across competitions	-0.00903	0.0234**	0.0338	0.0241
	(0.00988)	(0.00938)	(0.0239)	(0.0237)
Applicant is female	-0.171*	-0.0996	-0.129	-0.351***
	(0.0921)	(0.0937)	(0.122)	(0.130)
Applicant is from the Netherlands	0.102	0.145	0.126	0.138
	(0.213)	(0.139)	(0.167)	(0.176)
Applicant's year of birth	0.0166	0.0239**	0.0339**	0.0279*
	(0.0112)	(0.0114)	(0.0141)	(0.0149)
MNCS from 2 to 0 years before the application	0.181***	0.214***	0.230***	0.254***
	(0.0404)	(0.0414)	(0.0437)	(0.0527)
CWTS tool used information on applicant's e-mail address	0.292**	0.416***	0.410**	0.359*
	(0.144)	(0.144)	(0.196)	(0.209)
round	Yes	Yes	Yes	Yes
domain	Yes	Yes	Yes	Yes
Observations	992	803	539	427

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 22: MNJS smaller bandwidth regressions

	100%	75%	50%	25%
Received Veni grant ever	0.345**	-0.0459	-0.109	0.0184
	(0.167)	(0.150)	(0.238)	(0.242)
Rank normalized across competitions	-0.00360	0.0197***	0.0280	0.0168
	(0.00705)	(0.00710)	(0.0191)	(0.0185)
Applicant is female	-0.152**	-0.0782	-0.125	-0.296***
	(0.0721)	(0.0710)	(0.0978)	(0.101)
Applicant is from the Netherlands	-0.0268	0.0552	0.0237	0.0468
	(0.187)	(0.112)	(0.139)	(0.145)
Applicant's year of birth	0.0115	0.0147*	0.0185*	0.0125
	(0.00881)	(0.00860)	(0.0109)	(0.0116)
MNJS from 2 to 0 years before the application	0.178***	0.208***	0.232***	0.258***
	(0.0315)	(0.0313)	(0.0368)	(0.0367)
CWTS tool used information on applicant's e-mail address	0.276***	0.364***	0.379***	0.349**
	(0.104)	(0.0994)	(0.133)	(0.142)
round	Yes	Yes	Yes	Yes
domain	Yes	Yes	Yes	Yes
Observations	992	803	539	427

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: TNCS smaller bandwidth regressions

	100%	75%	50%	25%
Received Veni grant ever	9.089**	0.296	2.851	5.343
	(4.424)	(5.050)	(8.161)	(8.693)
Rank normalized across competitions	-0.223	0.326	0.0766	0.00722
	(0.169)	(0.226)	(0.623)	(0.585)
Applicant is female	-4.939**	-4.298**	-6.541**	-10.37***
	(2.061)	(2.072)	(2.857)	(3.495)
Applicant is from the Netherlands	5.591*	6.296**	5.532	7.420*
	(3.099)	(3.118)	(4.037)	(3.996)
Applicant's year of birth	0.355	0.419*	0.581*	0.526
	(0.238)	(0.254)	(0.344)	(0.352)
TNCS from 2 to 0 years before the application	1.323***	1.379***	1.398***	1.490***
	(0.244)	(0.254)	(0.286)	(0.304)
CWTS tool used information on applicant's e-mail address	2.591	8.517***	6.346	6.223
	(4.660)	(3.115)	(4.255)	(4.822)
round	Yes	Yes	Yes	Yes
domain	Yes	Yes	Yes	Yes
Observations	992	803	539	427

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: TNJS smaller bandwidth regressions

	100%	75%	50%	25%
Received Veni grant ever	7.228**	0.202	2.224	4.301
	(3.537)	(4.098)	(6.586)	(6.897)
Rank normalized across competitions	-0.195	0.244	0.0408	-0.136
	(0.132)	(0.184)	(0.500)	(0.455)
Applicant is female	-3.628**	-2.596	-4.146*	-7.020**
	(1.753)	(1.736)	(2.379)	(2.758)
Applicant is from the Netherlands	4.784*	5.330**	4.572	6.499**
	(2.602)	(2.535)	(3.278)	(3.295)
Applicant's year of birth	0.233	0.223	0.258	0.250
	(0.190)	(0.197)	(0.266)	(0.269)
TNJS from 2 to 0 years before the application	1.413***	1.429***	1.504***	1.573***
	(0.137)	(0.142)	(0.151)	(0.156)
CWTS tool used information on applicant's e-mail address	2.389	7.585***	6.004*	5.989
	(4.117)	(2.475)	(3.396)	(3.911)
round	Yes	Yes	Yes	Yes
domain	Yes	Yes	Yes	Yes
Observations	992	803	539	427

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

9.5 2SLS estimation with the priority score as the running variable

Just as in Section 6.3, I did the 2sls estimation with the four productivity indicators as dependent variables and six model specifications, the only difference being that I now used the priority score as the running variable.

For this analysis, I created a new estimation sample in the following way. After the removal of duplicates which I described in Section 5.1, my dataset contained 1,945 observations for which there was a priority score, a domain and a round. From these observations, 1,661 were first time applications. After collecting the CWTS output data, information on productivity was missing for 453 of these first time applications, leaving 1,202 observations for analysis. When removing 32 applicants who received the grant after 2008, as I did in the analysis with the rank, 1,180 observations were left to do analysis with.

The standard errors from the estimates from the analyses using the priority score as the running variable (shown in Tables 25-34) behave in similar ways as the standard errors from the estimates from the analysis with the rank in Section 6.3. They decrease with the inclusion of more controls and start increasing again when higher degree polynomials are added to the regression.

Just like in the 2sls estimations using the rank as the running variable, none of the grant coefficients are statistically significant at the 10% level for the short-term.

The long-term model (4) grant estimates are significant at the 10% level for all productivity indicators except publications, which has a grant estimate significant at the 5% level.

Unlike the grant estimates in my main analysis with the rank, the long-term model (5) grant estimates are statistically significant at the 10% level for all productivity indicators except the TNJS. This higher significance comes about because there are more observations to do analysis with.

For all productivity indicators, the estimated effect of the grant is higher when using the score than when using the rank as a running variable.

Table 25: Publications over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	-0.302	0.184	1.360	0.505	0.319	0.596
	(1.935)	(1.826)	(1.796)	(1.177)	(1.200)	(1.498)
Score normalized across competitions	-1.028	-1.015	-0.961	0.0444	-0.0761	0.100
	(0.684)	(0.684)	(0.656)	(0.421)	(0.462)	(0.680)
CWTS tool used information on applicant's e-mail address	4.625***	4.416***	3.657***	1.889***	1.903***	1.854***
	(0.897)	(0.895)	(0.829)	(0.652)	(0.646)	(0.645)
Applicant is female			-2.622***	-0.598	-0.589	-0.603
			(0.548)	(0.439)	(0.439)	(0.439)
Applicant's year of birth			0.110	0.0202	0.0229	0.0195
			(0.0802)	(0.0572)	(0.0563)	(0.0603)
Applicant is from the Netherlands				0.734	0.735	0.742
				(0.588)	(0.588)	(0.586)
Publications from 2 to 0 years before the application				1.769***	1.770***	1.769***
				(0.127)	(0.126)	(0.127)
(Normalized priority score)ˆ2					0.0981	0.119
					(0.128)	(0.154)
(Normalized priority score)ˆ3						-0.0260
						(0.0789)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 26: MNCS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.0941	0.0745	0.168	0.111	0.119	0.108
	(0.215)	(0.204)	(0.213)	(0.190)	(0.190)	(0.270)
Score normalized across competitions	-0.0738	-0.138**	-0.110	-0.0281	-0.0230	-0.0300
	(0.0683)	(0.0656)	(0.0669)	(0.0601)	(0.0616)	(0.117)
CWTS tool used information on applicant's e-mail address	0.0323	0.0366	-0.0464	-0.0226	-0.0233	-0.0213
	(0.190)	(0.181)	(0.189)	(0.163)	(0.163)	(0.171)
Applicant is female			-0.219***	-0.119*	-0.120*	-0.119*
			(0.0782)	(0.0688)	(0.0688)	(0.0691)
Applicant's year of birth			0.0200**	0.0123	0.0122	0.0124
			(0.00932)	(0.00839)	(0.00840)	(0.00871)
Applicant is from the Netherlands				-0.0217	-0.0218	-0.0220
				(0.102)	(0.102)	(0.102)
MNCS from 2 to 0 years before the application				0.384***	0.384***	0.384***
				(0.0420)	(0.0419)	(0.0421)
(Normalized priority score)ˆ2					-0.00418	-0.00499
					(0.0143)	(0.0168)
(Normalized priority score)ˆ3						0.00103
						(0.0108)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 27: MNJS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.121	0.114	0.194	0.0917	0.117	0.168
	(0.163)	(0.153)	(0.159)	(0.142)	(0.143)	(0.199)
Score normalized across competitions	-0.0672	-0.107**	-0.0849*	-0.0346	-0.0187	0.0135
	(0.0522)	(0.0495)	(0.0499)	(0.0446)	(0.0472)	(0.0855)
CWTS tool used information on applicant's e-mail address	0.118	0.0973	0.0301	0.0472	0.0450	0.0357
	(0.106)	(0.104)	(0.106)	(0.0936)	(0.0936)	(0.0985)
Applicant is female			-0.214***	-0.123**	-0.124**	-0.127**
			(0.0568)	(0.0512)	(0.0513)	(0.0518)
Applicant's year of birth			0.0225***	0.0140**	0.0136**	0.0131**
			(0.00643)	(0.00564)	(0.00565)	(0.00576)
Applicant is from the Netherlands				-0.0610	-0.0612	-0.0601
				(0.0876)	(0.0878)	(0.0876)
MNJS from 2 to 0 years before the application				0.358***	0.356***	0.355***
				(0.0323)	(0.0321)	(0.0324)
(Normalized priority score)ˆ2					-0.0131	-0.00941
					(0.0111)	(0.0128)
(Normalized priority score)ˆ3						-0.00475
						(0.00789)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 28: TNCS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	1.772	1.864	4.158	2.228	2.145	2.856
	(3.995)	(3.719)	(3.793)	(3.011)	(3.039)	(4.200)
Score normalized across competitions	-1.251	-1.723	-1.783	0.396	0.342	0.794
	(1.266)	(1.267)	(1.220)	(1.065)	(1.110)	(1.934)
CWTS tool used information on applicant's e-mail address	5.960***	6.444***	4.883***	3.238**	3.245**	3.120*
	(1.977)	(1.831)	(1.793)	(1.629)	(1.628)	(1.698)
Applicant is female			-5.004***	-2.171**	-2.167**	-2.202**
			(1.186)	(1.014)	(1.011)	(1.022)
Applicant's year of birth			0.245*	0.151	0.152	0.143
			(0.146)	(0.118)	(0.119)	(0.124)
Applicant is from the Netherlands				0.351	0.351	0.370
				(1.713)	(1.712)	(1.700)
TNCS from 2 to 0 years before the application				1.465***	1.465***	1.464***
				(0.184)	(0.184)	(0.184)
(Normalized priority score)ˆ2					0.0440	0.0965
					(0.241)	(0.299)
(Normalized priority score)ˆ3						-0.0667
						(0.170)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 29: TNJS over a period of 4 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	1.175	1.596	3.428	0.746	0.787	1.236
	(3.187)	(2.995)	(2.988)	(2.254)	(2.306)	(3.033)
Score normalized across competitions	-1.307	-1.475	-1.550	0.0984	0.125	0.409
	(1.056)	(1.059)	(0.993)	(0.754)	(0.815)	(1.342)
CWTS tool used information on applicant's e-mail address	5.934***	5.889***	4.512***	2.988***	2.985***	2.907**
	(1.374)	(1.442)	(1.345)	(1.100)	(1.100)	(1.148)
Applicant is female			-4.363***	-1.227	-1.229	-1.252
			(0.953)	(0.797)	(0.796)	(0.802)
Applicant's year of birth			0.263**	0.107	0.107	0.101
			(0.116)	(0.0906)	(0.0906)	(0.0941)
Applicant is from the Netherlands				0.420	0.420	0.430
				(1.186)	(1.187)	(1.181)
TNJS from 2 to 0 years before the application				1.552***	1.552***	1.551***
				(0.159)	(0.159)	(0.159)
(Normalized priority score)ˆ2					-0.0215	0.0116
					(0.198)	(0.238)
(Normalized priority score)ˆ3						-0.0420
						(0.124)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 30: Publications over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	3.261	4.390	6.660**	5.557**	5.191*	6.153
	(3.498)	(3.189)	(3.092)	(2.819)	(2.930)	(4.084)
Score normalized across competitions	-0.967	-0.302	0.182	1.059	0.822	1.434
	(1.134)	(1.077)	(1.018)	(0.898)	(0.999)	(1.772)
CWTS tool used information on applicant's e-mail address	7.061***	6.544***	5.179***	3.660**	3.688**	3.520**
	(1.561)	(1.536)	(1.531)	(1.553)	(1.533)	(1.477)
Applicant is female			-3.528***	-1.658*	-1.640*	-1.689*
			(0.945)	(0.887)	(0.888)	(0.912)
Applicant's year of birth			0.0995	0.0307	0.0360	0.0244
			(0.134)	(0.123)	(0.122)	(0.134)
Applicant is from the Netherlands				2.605**	2.607**	2.630**
				(1.212)	(1.212)	(1.221)
Publications from 2 to 0 years before the application				1.642***	1.644***	1.640***
				(0.111)	(0.112)	(0.115)
(Normalized priority score)ˆ2					0.193	0.264
					(0.253)	(0.281)
(Normalized priority score)ˆ3						-0.0902
						(0.183)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 31: MNCS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.392	0.380	0.569**	0.513*	0.424*	-0.0246
	(0.284)	(0.273)	(0.273)	(0.279)	(0.244)	(0.368)
Score normalized across competitions	0.0696	0.0334	0.0898	0.109	0.0521	-0.232
	(0.109)	(0.106)	(0.108)	(0.109)	(0.0855)	(0.200)
CWTS tool used information on applicant's e-mail address	0.448***	0.402***	0.306**	0.332**	0.340**	0.420***
	(0.148)	(0.135)	(0.138)	(0.140)	(0.138)	(0.148)
Applicant is female			-0.227***	-0.191**	-0.186**	-0.164*
			(0.0849)	(0.0855)	(0.0844)	(0.0859)
Applicant's year of birth			0.0212**	0.0200*	0.0213**	0.0266**
			(0.0102)	(0.0104)	(0.0104)	(0.0108)
Applicant is from the Netherlands				0.0125	0.0128	0.00117
				(0.209)	(0.209)	(0.208)
MNCS from 2 to 0 years before the application				0.143***	0.147***	0.155***
				(0.0319)	(0.0317)	(0.0318)
(Normalized priority score)ˆ2					0.0471	0.0142
					(0.0357)	(0.0275)
(Normalized priority score)ˆ3						0.0419
						(0.0278)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 32: MNJS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	0.307	0.309*	0.433**	0.363*	0.315*	0.167
	(0.198)	(0.183)	(0.184)	(0.194)	(0.185)	(0.266)
Score normalized across competitions	0.0271	0.00707	0.0470	0.0605	0.0302	-0.0630
	(0.0732)	(0.0682)	(0.0690)	(0.0705)	(0.0649)	(0.123)
CWTS tool used information on applicant's e-mail address	0.406***	0.355***	0.284***	0.308***	0.312***	0.339***
	(0.110)	(0.101)	(0.101)	(0.102)	(0.101)	(0.105)
Applicant is female			-0.196***	-0.154**	-0.151**	-0.143**
			(0.0658)	(0.0673)	(0.0667)	(0.0684)
Applicant's year of birth			0.0172**	0.0143*	0.0149*	0.0166**
			(0.00768)	(0.00778)	(0.00775)	(0.00813)
Applicant is from the Netherlands				-0.0572	-0.0567	-0.0601
				(0.180)	(0.180)	(0.179)
MNJS from 2 to 0 years before the application				0.161***	0.164***	0.169***
				(0.0312)	(0.0309)	(0.0317)
(Normalized priority score)ˆ2					0.0249	0.0141
					(0.0183)	(0.0208)
(Normalized priority score)ˆ3						0.0138
						(0.0123)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 33: TNCS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	9.311	10.80	14.81**	12.36*	10.86*	8.456
	(7.863)	(7.377)	(6.874)	(6.555)	(6.377)	(8.509)
Score normalized across competitions	0.175	1.114	1.478	3.247	2.281	0.752
	(2.594)	(2.563)	(2.405)	(2.350)	(2.231)	(3.877)
CWTS tool used information on applicant's e-mail address	9.328**	8.500**	4.778	3.243	3.357	3.781
	(4.396)	(4.272)	(4.318)	(4.590)	(4.491)	(4.228)
Applicant is female			-7.466***	-4.957**	-4.880**	-4.759**
			(2.127)	(2.107)	(2.090)	(2.108)
Applicant's year of birth			0.337	0.284	0.305	0.334
			(0.262)	(0.252)	(0.254)	(0.274)
Applicant is from the Netherlands				3.553	3.557	3.494
				(3.217)	(3.201)	(3.197)
TNCS from 2 to 0 years before the application				1.295***	1.301***	1.305***
				(0.233)	(0.233)	(0.233)
(Normalized priority score)ˆ2					0.790	0.612
					(0.703)	(0.640)
(Normalized priority score)ˆ3						0.225
						(0.459)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 34: TNJS over a period of 5 to 8 years after the grant application

	(1)	(2)	(3)	(4)	(5)	(6)
Received Veni grant ever	7.438	9.376	12.63**	9.602*	8.068	7.224
	(6.605)	(6.149)	(5.695)	(5.458)	(5.327)	(7.010)
Score normalized across competitions	-0.267	0.958	1.363	2.710	1.724	1.189
	(2.228)	(2.172)	(2.035)	(1.943)	(1.859)	(3.182)
CWTS tool used information on applicant's e-mail address	8.811**	7.560**	4.442	3.072	3.190	3.337
	(3.846)	(3.800)	(3.839)	(4.055)	(3.950)	(3.651)
Applicant is female			-6.257***	-3.430*	-3.350*	-3.307*
			(1.851)	(1.789)	(1.770)	(1.780)
Applicant's year of birth			0.319	0.204	0.226	0.236
			(0.215)	(0.201)	(0.202)	(0.220)
Applicant is from the Netherlands				3.520	3.530	3.511
				(2.636)	(2.618)	(2.612)
TNJS from 2 to 0 years before the application				1.393***	1.399***	1.401***
				(0.133)	(0.133)	(0.135)
(Normalized priority score)ˆ2					0.805	0.742
					(0.607)	(0.560)
(Normalized priority score)ˆ3						0.0789
						(0.392)
round	No	Yes	Yes	Yes	Yes	Yes
domain	No	Yes	Yes	Yes	Yes	Yes
Observations	1180	1180	1059	1046	1046	1046

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$