BGPE Discussion Paper

No. 124

The Effect of University Openings on Local Human Capital Formation: Difference-in-Differences Evidence from Germany

Benedikt Siegler

August 2012

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.
Friedrich-Alexander-University Erlangen-Nuremberg

B G P E Discussion Paper Series

The Effect of University Openings on Local Human Capital Formation: Difference-in-Differences Evidence from Germany

Benedikt Siegler[*]

August 2012

Abstract

Between 1960 and 1979, 93 new universities opened in Germany. Using this large tertiary education expansion, I estimate the effect of a university opening on the probability of obtaining a university degree in the local population. I exploit the geographical variation in local university access in a difference-in-differences approach by comparing age cohorts in counties that were and were not affected by the opening. Results show that a new university increases the share of university graduates in a county by 8 to 10 percentage points. The effect seems to be mainly driven my females and immigrants.

JEL Classification: I23, I28, H75

Keywords: tertiary education, education expansion, difference-in-differences, natural experiment, new university opening

Benedikt Siegler
University of Munich and Ifo Institute
Poschingerstraße 5
81679 Munich, Germany
+49 (0)89 9224 1240
siegler@ifo.de

[*] I thank Oliver Falck, Guido Schwerdt, Ludger Woessmann, and participants of the 12th BGPE Research Workshop for valuable comments. Financial support from the Bavarian Graduate Program in Economics (BGPE) is gratefully acknowledged. All remaining errors are my own.

1 Introduction

Increasing participation in tertiary education is a key policy objective in much of the world[1]. One way of increasing human capital investment at the tertiary level might be to extend the regional coverage of tertiary education opportunities by opening new universities in regions without prior local university access. In theory, the opening of a new university reduces the average costs of investment in tertiary education for the individual, leading to higher demand for tertiary education in the local population. However, it remains an empirical question to what extend this strategy actually works.

This paper investigates the importance of local university access to tertiary education participation by the local population. I collected information on new university openings during the large tertiary education expansion in Germany that occurred between 1960 and 1979, which I then linked to micro data on educational attainment from the German Socio-Economic Panel (SOEP). Exploiting variation in local university availability over time and between counties, I estimate the effect of local university access on the probability of obtaining a university degree in a difference-in-differences framework.

To my knowledge, this study is the first to evaluate the effect of a new university opening on tertiary education participation in the local population in Germany. Results show that the effect of a new university opening on the probability of obtaining a tertiary degree is between 8 and 10 percentage points for openings from 1960 through 1979. This is the effect in the local population including individuals without a university entrance certificate (Abitur). In aggregate terms, this means that the local rate of university graduates rises by 8 to 10 percentage points when a new university is established.

My results are robust to various model specifications that involve extensive controlling for confounding effects. The biggest effects are found for university openings in

[1] For example, the EU 2020 Agenda of the European Commission states the goal to raise tertiary education levels in member states to at least 40% among 30-34-year-olds (cf. http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/targets/index_en.htm).

counties without any previous tertiary education institution; the effect is lower for university openings in counties with previously existing tertiary education institutions.

In addition, I find that the effect is higher for females compared to males. It is also higher for immigrants compared to natives. Parental education background does not seem to matter much, although individuals with low-educated parents seem to profit somewhat more.

The remainder of the paper is organized as follows. Section 2 presents an overview of related literature. In Section 3, I briefly describe the process of the tertiary education expansion in Germany after WW II. A description of the data and the methodological framework is provided in Section 4. The main regression results are presented in Section 5. In Section 6, I discuss some issues in regard to the identification strategy and show several robustness checks. Section 7 concludes.

2 Related Literature

Previous studies identify parental education as a relevant determinant of the decision to enroll in tertiary education (e.g., Cameron and Heckman, 2001; Burnhill et al., 1990). In a recent study for Germany, however, Riphahn and Schieferdecker (2012) analyze intergenerational education mobility and find that in the transition to tertiary education, the effect of parental education, though important, is less significant than the effect of parental income. It makes intuitive sense that financial constraints could inhibit the transition to tertiary education. As shown by Dynarski (2003), student financial aid has a strong positive effect on college attendance and completion rates. She finds that a $1,000 student benefit increases the share of high school graduates enrolled in higher education by 3.6 percentage points. Other studies that focus on the United States find similar results (e.g., Leslie and Brinkman, 1988; Kane, 1994; Dynarski, 2000). In a recent study for Germany, Steiner and Wrohlich (2011) find a positive

but much lower effect of public student aid on enrollment rates. They calculate a 1.5 percent increase per €1,000 student aid.

Being located far from a university can pose considerable cost of obtaining a tertiary education. In this respect, the local availability of a university decreases the access costs and thus may act as a strong incentive for investment in tertiary education. Several studies relying on this distance-cost argument exploit the variation in local university access in order to obtain consistent estimates of the returns to education (e.g., Card, 1995; Currie and Moretti, 2003; Moretti, 2004). The relationship between access costs and enrollment rates is further explored in two recent studies, both of which found large and robust effects (Frenette, 2009, for Canada; Spiess and Wrohlich, 2010, for Germany). Frenette (2009) examines the effect of a local university on tertiary enrollment rates in Canada. In his empirical approach, he uses a dummy variable indicating the presence of a local university in a census metropolitan area to estimate the effect of distance on the probability of attending tertiary education. The results indicate a 6.4 percentage point increase in the university attendance rate when a local university is available. However, this result is mainly due to a substitution effect from college to university. The net effect is a 1.3 percentage point increase. Spiess and Wrohlich (2010) analyze the enrollment decisions of German high school graduates conditional on the distance to the nearest university. Their estimation results suggest that there is a threshold around 12.5 km. Individuals at a distance above this threshold have a 7 percentage points lower probability of attending a university compared to those living within a range of 6.5 km from the nearest university. Below this threshold of 12.5 km, distance does not seem to have an effect on university attendance. This is an interesting, but hard to explain, result, implying as it does that a mere increase of 6 km has a huge inhibiting effect on university attendance. Still, the overall result is that distance imposes costs and thus has an inhibiting effect on university enrollment.

3 A Brief Overview of Tertiary Education Expansion in Germany

Expansion of the tertiary education sector in Germany after World War II occurred in three main phases. The first increase in the number of universities occurred during the post–World-War-II era. The second and largest increase took place during the 1960s and 1970s. It was during this period that a new type of university was introduced into the higher education system, the so-called universities of applied sciences (*Fachhochschulen*). The third large expansion of tertiary education institutions happened in the 1990s. However, a significant part of this last increase was due to the inclusion of universities from East Germany after the reunification in 1990. Figure 1 illustrates this expansion process.

<<Figure 1 about here>>

Figure 2 shows the number of university foundations per year and provides a clear illustration of how massive the expansion activity was, particularly between 1960 and 1979. The expansion peaked in 1971 with the foundation of 42 new universities.[2]

<<Figure 2 about here>>

Universities of applied sciences (*Fachhochschulen*) were first introduced by state law in 1968. They were intended to provide a higher education with a strong focus on the practical application of the concepts taught, including a mandatory internship at a firm. Indeed, a key characteristic of a university of applied sciences is its strong cooperation with local firms. Thus, the needs and requirements of the local labor market are taken into account in the educational qualification process. The range of subjects taught by these universities has

[2] Some universities (especially universities of applied sciences) have several divisions, which are sometimes located up to 50 km apart from each other in different towns. When constructing the graph, these divisions were treated as independent entities.

increased substantially since the early days, but the focus is still on technical and social disciplines. This expansion of the tertiary education system allowed more students to enroll in university studies. The rapid increase in the number of students is shown in Figure 3.

<<Figure 3 about here>>

There were many reasons behind expanding the tertiary education sector. At the beginning of the process in the 1950s, one major aspect was the Cold War. Supporters of the expansion argued that without massive investment in human capital, the German economy would fail to effectively compete with the Soviet Union.[3] Over time, however, the founding of a new university was accompanied by less political ideology and the reasons put forth began to have more to do with stimulating the local economy of economically underdeveloped regions. In order to rapidly increase the supply of tertiary education, the majority of new universities founded in the 1970s used the infrastructure of previously existing institutions such as schools of engineering. These institutions, however, were neither homogeneously structured nor part of the tertiary education system. Since the 1990s, there have been increasingly more new foundations in places that did not previously have an institution of higher education. Figure 4 shows the geographic distribution of universities in Germany over time.

<<Figure 4 about here>>

4 Estimation Strategy

I use data from the German Socio-Economic Panel (SOEP), a large panel study that has been running since 1984 with annual follow-up. In 2009, more than 20,000 individuals in roughly

[3] Cf. Picht (1964).

12,000 households were interviewed. The SOEP is a representative survey of the population living in Germany and contains rich information on socioeconomic variables. Despite all these advantages, however, I cannot analyze the majority of university openings using the panel structure of the SOEP.[4] Therefore, I use only the 2009 wave. For every individual, I have information on the highest degree achieved. I also know the county of residence of the individuals in 2009.[5] In addition to the SOEP data, I collected comprehensive information about the universities and universities of applied sciences in Germany, including the date and place of their foundation.[6] In a second step, I merged this information with the SOEP data at the county level. This means that all individuals who lived in a particular county in 2009 were assigned the same information regarding university openings.

The aim of this study is to estimate the effect of a new university opening on the probability of obtaining a tertiary degree in the local population.[7] In an ideal experiment, the first step would be to randomly assign a pool of individuals to two groups. Individuals in the treatment group would obtain access to a local university; individuals in the control group would not. Then, after a certain passage of time, one would compare the two groups with respect to the share of university graduates. However, aside from the general problem that such an experiment is not feasible, other concerns arise. For example, how long should one wait before comparing outcomes? Moreover, not only does assignment to a group have to be random, but so does the selection of individuals with respect to the areas in which they grew

[4] This is because the majority of new university openings occurred in the 1960s and 1970s, prior to initiation of the SOEP.

[5] Due to privacy protection when using regional identifiers, all data processing and subsequent analysis was done using remote access to the data.

[6] A list of universities in Germany, including detailed information, can be downloaded from the webpage of the German university association at: http://www.hochschulkompass.de/index.html. However, the information supplied is not always accurate. To ensure that my dates were correct, I went through each university's individual webpage to verify its date of opening.

[7] In contrast to many existing studies, I consider university completion rather than university enrollment as the outcome variable, for two reasons. First, a university degree is an important job market certificate. Second, the difference between the number of students in the first year of university studies and those who actually complete a degree can be substantial; based on administrative data for 2006, the student drop-out rate at German universities is about 30 percent (cf. Heublein et al., 2008).

up. This is important because the treatment effect might differ between those individuals who grew up near a university and those with no university nearby.

I solve this problem by exploiting the tertiary education expansion of the 1960s and 1970s in Germany, which can be viewed as a natural experiment. Within this 20-year period (1960 through 1979), 93 new universities were opened. Forty-four of these new universities were established in regions in which there previously was no other tertiary education institution (henceforth called "green-field university openings"). I use an individual's age to relate him or her to a particular university opening. For any county where a new university was opened, I can then compare those individuals who still needed to decide whether to attend university with those individuals for whom it is reasonable to assume that they had already made this decision (at the time of the new university opening). The key identifying assumption is that career plans are made at a certain age. Usually, the decision to attend university is made at the time one receives the university entrance diploma (Abitur) (cf. Spieß and Wrohlich, 2010, p. 473), typically around the age of 19 or 20. Using a simple cohort comparison to identify the effect from university opening, however, would likely suffer a bias from a general trend toward more tertiary education over time. To disentangle the treatment effect of the university opening from any underlying time effects, I use a treatment-control framework in the sense of a difference-in-differences (DiD) estimator. When a new university opened, individuals living in its proximity experienced a sharp decrease in access costs, whereas for individuals living far away this decrease can be expected to be negligible. For the sake of feasibility, I use counties to distinguish between treatment and control. In this sense, the treatment group consists of individuals living in counties where a university was opened between 1960 and 1979. The control group is comprised of those individuals living in counties that did not have any tertiary education institution at that time.

But what exactly is the DiD estimator capable of identifying in this framework? Consider two counties, A and B. County A acquires a new university, county B does not;

thus, I consider A a treatment county and B a control county. Under the assumption that the university opening only affects the population of county A, I can calculate the exact treatment effect of this opening. However, the university opening in A likely affects individuals in B as well. If this effect is positive, individuals may move or commute to the university in county A. Econometrically, this would lead to an underestimation of the treatment effect. The result would be a lower-bound estimate. However, at least theoretically, there is also the possibility that the effect of a new university opening in county A can have a negative effect in county B, as a result of a general equilibrium effect. Due to increased tertiary education enrollment in county A, the return expectations from tertiary education might fall, so that in county B, where access costs are still higher, fewer individuals than before obtain a university degree. Consequently, the DiD estimator would overstate the true effect of a university opening in my framework. However, as I show later, commuting is common, and thus the estimates should be interpreted rather as lower bounds.

To construct the dummy variables for the DiD framework, I use the following technical procedure. In a first step, I assign each individual in my data set to either the treatment or the control group, depending on the residential information he or she provided in 2009. Individuals living in counties where a green-field university opened within the considered period were assigned to the treatment group. Individuals in counties without a tertiary education institution were assigned to the control group. In the second step, I apply a nesting strategy, such that all different openings can be analyzed simultaneously using the DiD approach. This step is necessary because, otherwise, cell sizes would be too small for statistical inference. For every university opening, I select those individuals who were 16–28 years of age. I do this separately for all new openings. For every university opening in the treatment group, I construct corresponding age groups in the control group (counterfactual treatments). I then pool all observations. This procedure implies that some observations of the control group occur more than once in the final data set. For example, if a person was age 16

when a university was opened and age 23 at the time of another university opening, the person shows up twice in the final data set. To account for the replication of observations in the subsequent analysis, I weight all observations with the inverse of their occurrence in the data set. To empirically distinguish the actual threshold for the cohort comparison, I plot the share of individuals with a university degree conditional on the age group they belonged to at the time a new university was established (see Figure 5).

<<Figure 5 about here>>

The group of individuals that live in counties without a university throughout the period of analysis (indicated by the blue line in Figure 5) has a fairly even distribution of university graduates (around 19 percent) across all age groups. For individuals living in counties where a new university was established, one can distinguish a clear difference in the share of university graduates between those aged 21 and younger and those aged 22 and older. For the older cohorts, the share of university graduates is at roughly the same level as in the control counties. For the younger cohorts, it is much higher, at 24 to 28.5 percent. This finding is consistent with the hypothesis that, in general, career plans are made at a certain age, which coincides with the usual university entrance age in Germany. At least from the data at hand there is no indication that those individuals older than 21 were also affected by a new university opening. A similar picture emerges when those counties that already had at least one university are also considered. This is demonstrated in Figure 6.

<<Figure 6 about here>>

The inclusion of counties with previously existing universities induces a level effect in the "treated"-counties share, whereas the "control"-counties share is relatively unchanged. Both

Figures 5 and 6 suggest that the threshold for the cohort comparison lies between 21 and 22 years of age.

To discover the effect of a university opening on the probability of obtaining tertiary education, I estimate the following equation using OLS:

$$(1) \quad University\_degree_{ijk} = \beta_0 + \beta_1 Uni\_opening_j + \beta_2 Agegroup_k + \beta_3 \left(Uni\_opening_j \times Agegroup_k\right) + X'_{ijk}\beta + \alpha_j + \gamma_k + u_{ijk}.$$

The outcome variable in Equation (1) is a dummy variable for whether an individual *i* obtained a university degree (= 1) or not (= 0). *Uni_opening* indicates a dummy variable that refers to the county *j* of residence and equals 1 for an individual who lives in a county where a new university was opened; 0 otherwise. *Agegroup* denotes a dummy variable that equals 1 if the individual is 16–21 years of age. It takes the value 0 for an individual 22–28 years of age. *X* is a vector of control variables, α indicates county fixed effects, γ represents cohort fixed effects, and u denotes an idiosyncratic error term. All regressions were performed using clustered standard errors. I constructed the clusters to account for the dependence of observations within groups of county and year of birth. In particular, I interacted the county of residence and the year of birth of an individual to obtain the cluster variable.

Table 1 shows descriptive statistics of variables used in the analysis. As shown in Column (1), roughly 20 percent of the individuals in my sample have a tertiary degree, slightly more than 50 percent are female, about 10 percent have at least one parent who graduated from university. The share of immigrants is about 8.5 percent. In Columns (2) and (3) these sample means are calculated separately for individuals living in counties with and without a university opening. Column (4) shows results of a test for difference of the means.

With the exception of the share of immigrants both groups (treatment and control) are balanced with respect to these key variables used in the regressions.

<<Table 1 about here>>

5 Results

5.1 Effect of University Opening on Obtaining Tertiary Degree

Table 2a shows the results from a linear regression using all green-field university openings between 1960 and 1979.[8] The dependent variable is a dummy variable indicating whether or not a person obtained a university degree. In the baseline specification (1), the only explanatory variables are the three dummy variables from the DiD framework (*Uni_opening*, reflecting whether the individual lived in the county of the particular university opening; *Agegroup*, indicating whether the individual was 16–21 years of age when the new university opened; and an interaction term of the two whose coefficient gives the treatment effect). In the other specifications (2)–(5), other explanatory variables are added, that is, the person's gender, an indicator variable for whether one of the parents obtained a tertiary degree, and whether or not the person is an immigrant. In specifications (4) and (5), I added cohort fixed effects as well as county fixed effects (Column (5)) to account for potential time- and region-specific differences that might bias the coefficient of interest.[9] Throughout all specifications, the results are extremely robust. The treatment effect ranges from 8 to 10 percentage points and is statistically significant at the 5 percent level (p-value ≤ 0.05). This means that the opening of a new university is associated with an 8 to 10 percentage point increase in university graduates among local youth.

[8] I also ran nonlinear (probit) regressions; results were similar.

[9] The effect also remained robust when I estimated a model with an additional linear age trend. The point estimate was a little higher compared to specification (5) (0.088), however significance went down.

<<Table 2a about here>>

The effect from a green-field university opening (i.e., going from no tertiary education capacity to at least some capacity) might be substantially different than the effect from additional openings. A change in the extensive margin of tertiary education supply gives students for whom distance imposes inhibitive costs the opportunity to pursue university studies, whereas a change in the intensive margin is usually associated with a greater variety of courses and thus should instead affect the choice of subject studied. Table 2b presents estimation results using all 93 university openings between 1960 and 1979. In addition to the 44 green-field university openings, there were 49 openings in counties where at least one university already existed. When these 49 additional openings are included in the sample, the estimated treatment effect is much lower, ranging from 4.4 to 6.1 percentage points. This is consistent with the hypothesis that an additional university opening has a much lower effect on the probability of participating in tertiary education than does the first opening in a county.

<<Table 2b about here>>

5.2 Heterogeneous effects by gender, migration and parental education background

Another interesting question is to what extend the treatment effect might differ between individuals with different social backgrounds. For this purpose, I ran separate regressions on subsamples with respect to gender, migration status and parental education.

Table 3a shows separate results for males and females. The treatment effect from a university opening for males is only 2.9 percentage points and not statistically different from zero. For females, however, this effect amounts to 14 percentage points and is highly statistically significant. This result suggests that at least back in the 1960s and 1970s, females

had higher costs of obtaining a tertiary degree than males. One explanation might be that the role of females at the time was seen primarily as becoming a housewife and mother, so that financial resources within a household were first allocated to the education of males. When access costs decreased with the opening of a new university, females were able to obtain tertiary education.

Table 3b shows separate regression results for immigrants and natives. It appears that the treatment effect for immigrants is roughly twice as large as for natives. The major influx of immigrants into Germany at the beginning of the 1960s was caused by the need for unskilled labor at the time of the economic miracle. An immigrant household thus had less financial resources than a typical German household. Again, it is plausible to assume that the opening of a university reduced the costs for immigrant households to send their children to university.

Table 3c shows results from separate regressions on subsamples of individuals with respect to the level of parental education. I classified individuals where at least one parent graduated from university as having high-educated parents. On the other hand if neither of an individual's parents graduated from university, I classified him or her as having low educated parents. Although the treatment effect is not statistically significant from zero in neither subgroup, it appears that the point estimate is slightly higher for individuals with low educated parents. Again it is intuitive to assume that households with low-educated parents have on average less financial resources and are thus constrained in sending their children to a faraway university.

Based on these results, it seems that females and immigrants have profited the most from the 1960s and 1970s tertiary education expansion in Germany. In general, these results provide additional support for the hypothesis that reducing access costs to tertiary education facilities can help increase the share of university graduates.

6 Robustness

6.1 Student Mobility: Relocation and Commuting

The data used in this study lack any information as to the place of residence of an individual at the time he or she went to university. Unfortunately, I cannot use the panel structure of the data set, either, because the greatest part of the tertiary education expansion took place in the 1960s and 1970s and the SOEP study only started in 1984. In this section, I first explain how unobserved student mobility can lead to biased estimates; I then present evidence for why it is reasonable to believe that student mobility is not a problem to the results in this study.

There are two major points in time when the issue of student mobility arises: (a) when students leave home to go to university in a different part of the country, and (b) after they complete their studies and move to their new workplace, if it is in a different region. The chief reason behind a student leaving home in order to attend a faraway university rather than a nearby institution should be the university's reputation.[10] If students move to those places that host a renowned university and stay there after completing their studies, they increase the share of educated people in those counties. The most renowned universities in Germany, however, were established a long time ago. Their incentive character should be largely independent from any new university foundation. The methodology I apply makes use of a cohort comparison. Cohorts below and above the typical university entrance age (at the time a new university was established) are compared with respect to their highest educational degree. An old university affects both these "before" and "after" cohorts to the same extent because it was established a long time ago. In this way, student mobility associated with old (potentially more attractive) universities only induces a level effect in the DiD framework and does not cause distortion when estimating the treatment effect. In a similar way, one can think about the question of whether student mobility after university completion is a problem for my

[10] For some subjects (like medicine), the allocation of places to students is centralized and lottery-like. These cases should not be a problem.

results. Mobility at this point in time should be mainly driven by job market considerations. Here, big urban centers offer a greater variety of employment. These places also offer higher living standards, another incentive for highly educated people (Falck et al., 2011). The following example should help clarify this point. Suppose that a university was opened in a county that had not previously had one. Motivated by this opening, some students of this county obtain tertiary education. However, after completing their studies, they move away to work somewhere else. In my retrospective analysis, these people are then missing from the share of university graduates in that county. The treatment effect of the university opening on graduation rates is then underestimated. An overestimation would occur if highly educated people from other counties moved to the county where the new university was opened. This would drive up the share of college graduates. In a subsequent analysis one would confound this mobility effect with the treatment effect of the university opening.

In the previous section, my baseline results were obtained using only green-field university openings between 1960 and 1979. This sample did not include counties with old (and potentially renowned) universities. It also did not include counties with big urban centers, as almost all bigger cities already had a university by that time. In this sense, I already might be controlling implicitly for student relocation as described above. Yet, a more direct measure is preferable. In another approach, I make use of a variable in the SOEP data set that allows me to control more directly for individuals' mobility. The variable is derived from a question asking individuals whether their current place of residence is the same as that of their childhood.[11] Restricting the sample to those people who never left their place of childhood enables me to evaluate the severity of the mobility concern. Table 4 shows

[11] There are two questions. The first question asks where a person lived until age 15. The second question asks whether the person is still living in the same place today. The possible choices for answering are: 1. Yes, still; 2. Yes, again; 3. No. Persons, who chose the first option, can be regarded as never having moved away from the area of their childhood. This implies that if this person obtained a university degree, he or she did so in that particular area. However, these questions are only asked once, when the person is interviewed for the first time. For the robustness analysis, I only use those individuals that in 2009 still lived in the same county as they did when entering the SOEP.

regression results using only those individuals who indicated never having left their place of childhood. Throughout all specifications, the effect of a new university opening is statistically significant and very robust, ranging from 12.7 to 14.6 percentage points. Compared to the estimates from Table 2a, the effect of a new university opening is about 4 percentage points higher. The explanation is intuitive: the most immobile individuals are those for whom the costs of moving are the highest for whatever reason. When a new university is opened in a place that did not previously have one, access costs are reduced and the impediment against going to university for the immobile individual is removed.

The reason I refrained from focusing entirely on immobile individuals is that in general inferences drawn from these individuals cannot be carried over to the "average" individual. In this sense, the results shown in Table 4 are local average treatment effects (LATEs).

<<Table 4 about here>>

As one would have expected the LATE estimates to be higher than on average, the results shown in Table 4 are evidence that the treatment effect from a new university opening in my primary specification is not overestimated due to student relocation. However, the effect can still be underestimated due to student commuting. The methodology I use to derive consistent estimates makes use of geographic variation. In particular, I use counties as a feasible spatial unit. In the DiD framework, these counties determine the separation of individuals into "treatment" and "control" groups. But a new university opening may also affect individuals in neighboring counties. If the distance is not too great (i.e., within a reasonable commute), people can commute between their county of residence and the county of the new university. In this way, even "immobile" individuals in counties belonging to the control group can obtain a tertiary education degree. Evidence for the existence of a

commuting effect is given in Table A1. It contains sample means on selected variables of immobile individuals by group (treatment vs. control). Column (2) shows these values for the group of immobile individuals in counties without any tertiary education institution. Without a commuting effect, the share of university graduates should be zero. However, 13.3 percent of these individuals indicated having a tertiary degree, which can be explained only by a commuting effect. This effect drives down the point estimates of the university opening effect. The obtained results should therefore be viewed as lower-bound estimates.

6.2 Treatment Assignment

The identification strategy relies on the assumption that the assignment of treatment, i.e. the opening of a new university, was random. One concern in this regard may be that the opening of new universities was based on demand considerations in the respective areas. If the new universities were systematically founded in areas with higher demand for tertiary education, the results would be biased upwards.

Although I cannot entirely rule out this possibility, there are several reasons to belief that the process of university openings was to a large extend not based on demand considerations. As described in Section 3, there were several reasons for the expansion of tertiary education. In the 1960s and 1970s, policymakers wanted to increase the supply of tertiary education as quickly as possible and thus used the preexisting infrastructure of schools of engineering to build new universities. This strategy placed considerable limitations on location options for new universities. Moreover, even after this period of expansion, decisions as to where to found a new university were driven by many things besides demand considerations. The process of public university foundation is complex and involves many different stakeholders. Like most other public investments in infrastructure, it needs to overcome several bureaucratic hurdles before a decision is made.

Additionally, I can compare the "pre-treatment" trends in the share of university graduates in the treatment and the control counties. Before the actual treatment (i.e., the opening of the new university), the trends in the share of university graduates should be similar. The shares of university graduates for both treatment and control groups are plotted in Figure 7.

<<Figure 7 about here>>

Figure 7 reveals different trends for cohorts aged 21 or younger, indicating the treatment effect of a new university opening. For older cohorts (i.e., individuals 22 years or older at the time of the university foundation) trends are fairly similar. Obviously, there must have been stronger determinants of new university foundation than merely demand. This finding supports the view that university foundation can be regarded as random with respect to demand and that the estimated treatment effect is not upward biased.

6.3 External validity

Before making policy recommendations, it is necessary to ensure that the estimated treatment effect can still be expected to hold nowadays. The results in this study are obtained by evaluating the effect of university openings in the 1960s and 1970s. Without further evidence, it is hard to argue that the magnitude of the effect would be the same today. In order to shed some light on this question, I ran regressions on a much broader sample of university openings. Table 5 shows regression results using university openings from 1946 to 1996. In Columns (1) through (5), only green-field university openings were considered, a total of 79 openings.

<<Table 5 about here>>

Compared to the results in Table 1a, the effects are somewhat lower. One explanation for this might be the inclusion of East German universities. After reunification, most of the existing East German universities were rearranged in terms of their administrative structure and as such are classified as new university openings in my sample, although this was not really the case. A separate investigation fails due to small sample sizes. When all 155 university openings between 1946 and 1996 are considered (Columns (6) through (10)), results are in the range of those from Table 1b. The inclusion of 62 additional university openings (mainly from the 1990s) did not seem to change the effect in any way. This suggests that we might still expect a quite substantial effect from a new university opening today.

7 Conclusion

This paper estimates the impact of a new university opening on the probability of obtaining a university degree in the local population by exploiting the large tertiary education expansion that occurred in Germany during the 1960s and 1970s. This expansion can be regarded as a natural experiment with respect to local university access. A substantial number of the new universities were built in regions that did not previously have a tertiary education institution. I use this exogenous variation in local university access to construct a difference-in-differences estimator and to calculate a treatment effect from university opening.

Results suggest that a new university opening in a place without any prior local access to tertiary education increases the probability of obtaining a university degree by 8 to 10 percentage points. The effect is lower for openings in regions with already existing tertiary education supply. The estimates are extremely robust. I also find that the effect of a new university opening was stronger among females than among males and twice as large for immigrants compared to natives. However, these results might reflect unique circumstances at that time, as these two groups were arguable more financially restricted.

The results have important policy implications. Policymakers may hope to promote investment in human capital at the tertiary level by opening new universities in regions without prior local university access. Theoretically, this strategy may work as it reduces the average cost of human capital investment at the tertiary level in the presence of mobility costs. The findings of this study provide empirical support that opening a new university does indeed increase participation in tertiary education.

References

Burnhill, P., Garner, C., McPherson, A. (1990). Parental Education, Social Class and Entry to Higher Education 1976–86. *Journal of the Royal Statistical Society*, 153 (2), 233–248.

Cameron, S. V., Heckman, J. J. (2001). The Dynamics of Educational Attainment for Black, Hispanic, and White Males. *Journal of Political Economy*, 109 (3), 455–499.

Card, D. (1995). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. In L. Christofides, E. K. Grant, R. Swidinsky (eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press, 201–222.

Currie, J., Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *Quarterly Journal of Economics*, 118 (4), 1495–1532.

Dynarski, S. (2000), Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance. *National Tax Journal*, 53 (3), 629–661.

Dynarski, S. (2003). Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion. *American Economic Review*, 93 (1), 279–288.

European Commission. (2012). Europe 2020 Targets. http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/targets/index_en.htm (Access: January 2012).

Falck, O., Fritsch, M., Heblich, S. (2011). The Phantom of the Opera: Cultural Amenities, Human Capital, and Regional Economic Growth. *Labour Economics*, 18 (6), 755–766.

Frenette, M. (2009). Do Universities Benefit Local Youth? Evidence from the Creation of New Universities. *Economics of Education Review*, 28 (3), 318–328.

Heublein, U., Schmelzer, R., Sommer, D. (2008). *Die Entwicklung der Studienabbruchquote an den deutschen Hochschulen*. HIS: Projektbericht.

HRK – Hochschulrektorenkonferenz. (2011). Hochschulkompass. http://www.hochschulkompass.de/ (Download: April 2011).

Kane, T. J. (1994). College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education. *Journal of Political Economy*, 102 (5), 878–911.

Leslie, L., Brinkman, P. (1988). *The Economic Value of Higher Education*. New York: Macmillan.

Moretti, E. (2004). Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data. *Journal of Econometrics*, 121, 175–212.

Picht, G. (1964). *Die deutsche Bildungskatastrophe, Analyse und Dokumentation*. Freiburg: Herder.

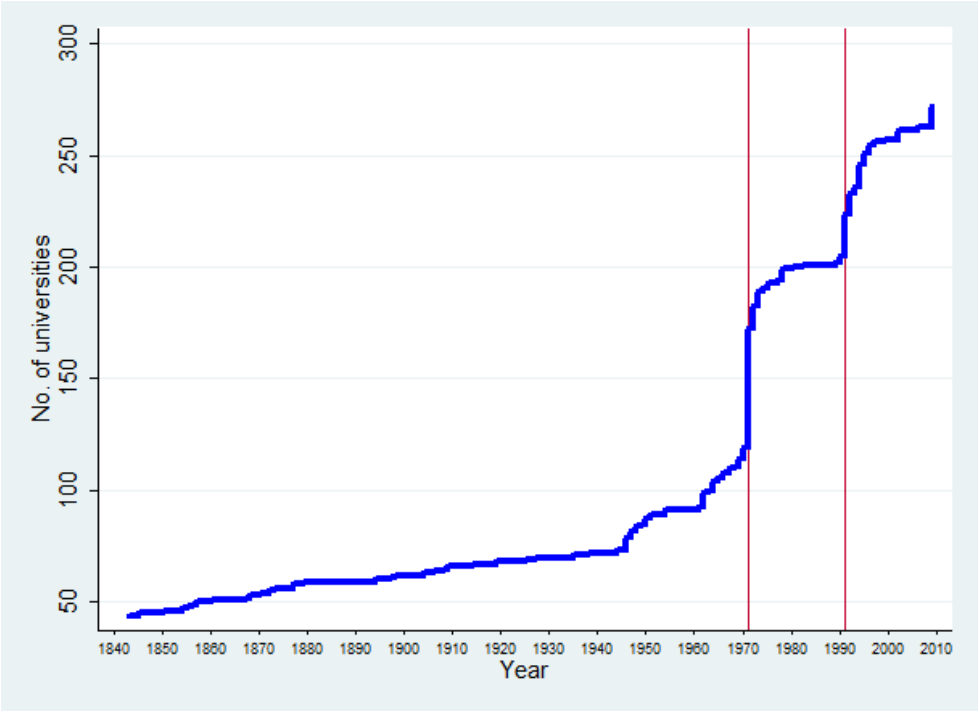
Riphahn, R., Schieferdecker, F. (2012). The Transition to Tertiary Education and Parental Background Over Time. *Journal of Population Economics*, 25 (2), 635–675.

Rüegg, W. (2004). *Geschichte der Universität in Europa: Vom 19. Jahrhundert zum zweiten Weltkrieg 1800–1945* (Vol. 3). München: C.H. Beck.

Spiess, K., Wrohlich, K. (2010). Does Distance Determine Who Attends a University in Germany? *Economics of Education Review*, 29 (3), 470–479.

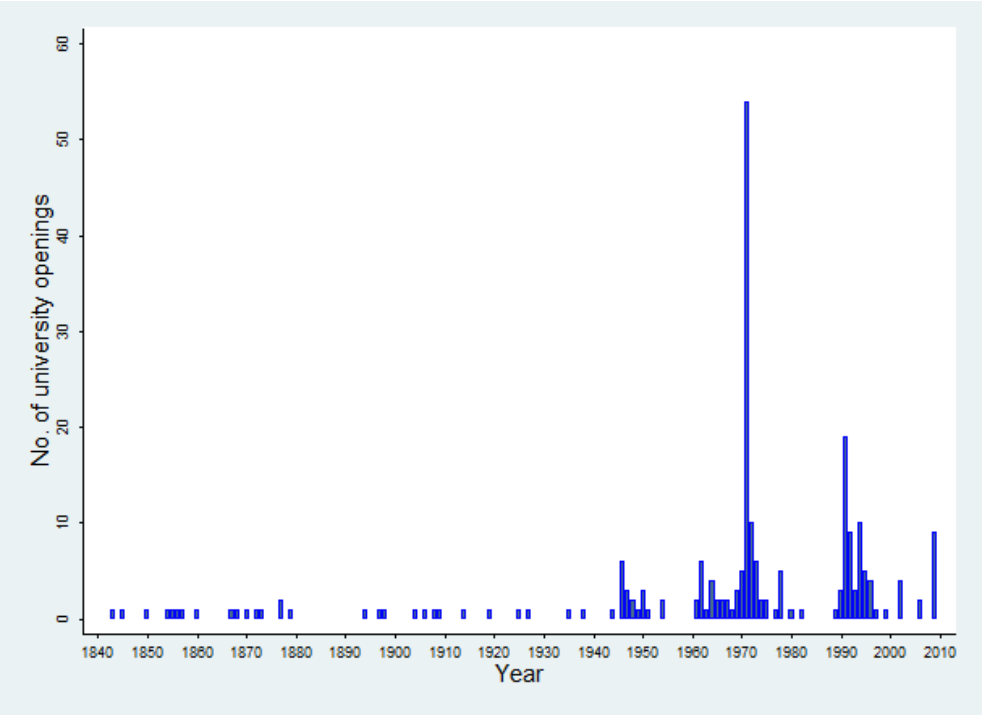
Steiner, V., Wrohlich, K. (2012). Financial Student Aid and Enrollment in Higher Education: New Evidence from Germany. *Scandinavian Journal of Economics*, 114 (1), 124–147.

Figure 1: Number of public universities in Germany



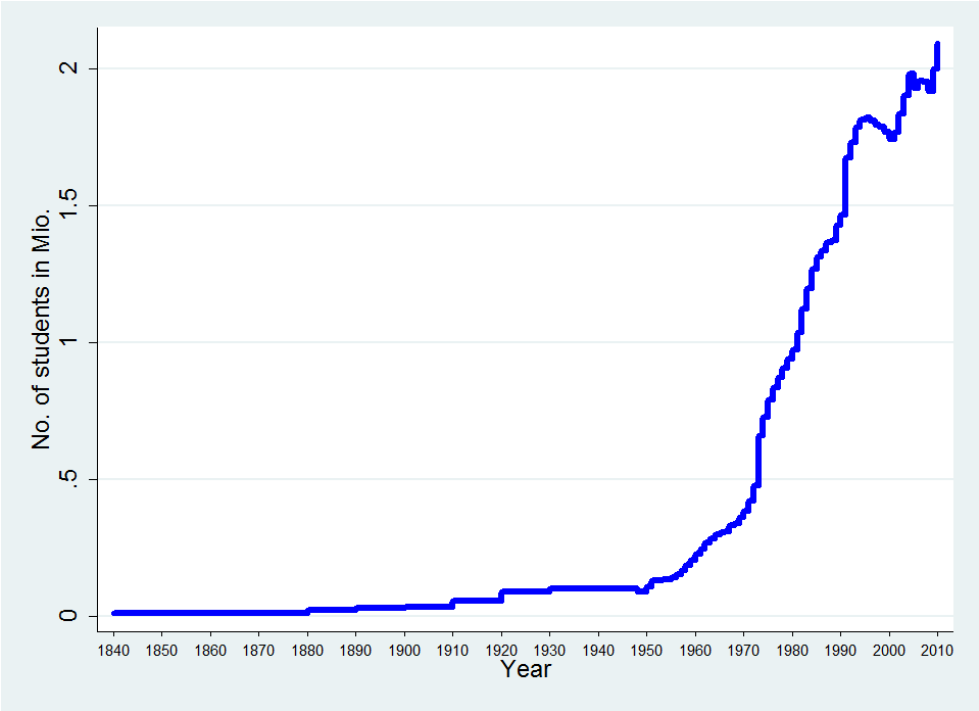
Notes: The graph shows the number of public universities in Germany. The first vertical red line indicates the first introduction of universities of applied sciences (Fachhochschulen) in 1971. The second vertical line indicates the first inclusion of former East German universities in 1991. Data from HRK (2011), own calculations.

Figure 2: Number of new university openings per year



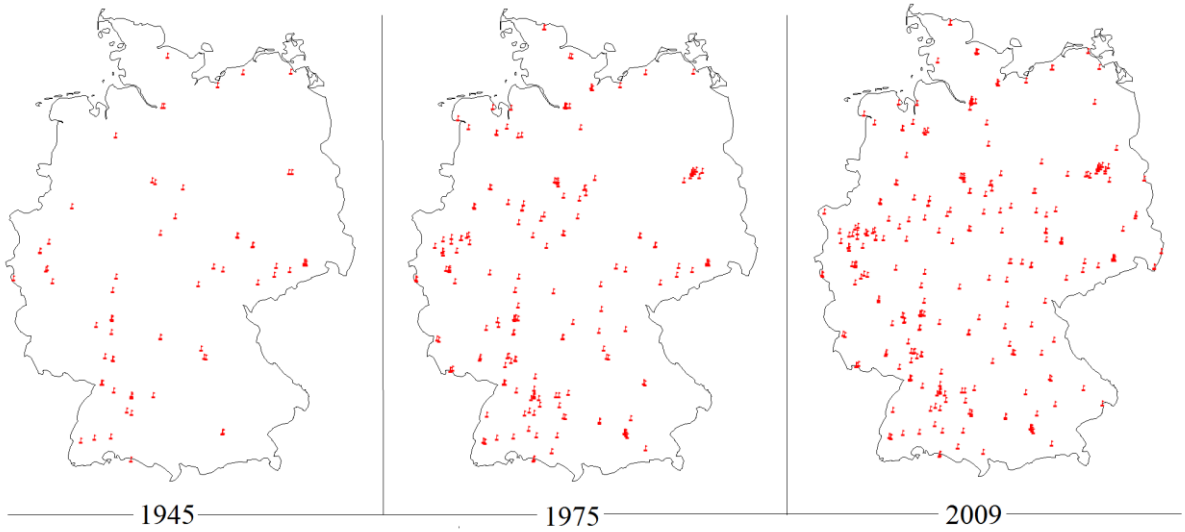
Notes: The graph shows the number of new university openings per year. When an institution had several divisions at different places, these were counted as independent entities. Data from HRS (2011), own calculations.

Figure 3: Number of students in Germany



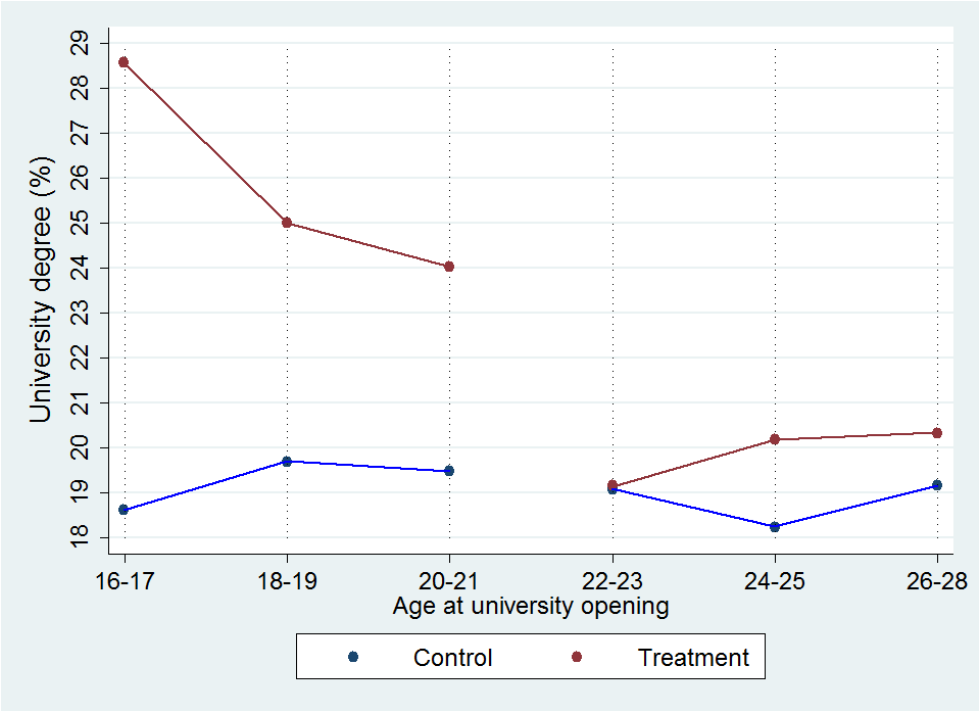
Notes: The graph shows the development in the number of students in Germany from the 1900s until 2010. From 1948 onward, the data are from official records supplied by the BMBF webpage. Data for years prior to 1948 are from Rüegg (2004). Own calculations.

Figure 4: Distribution of universities over space and time



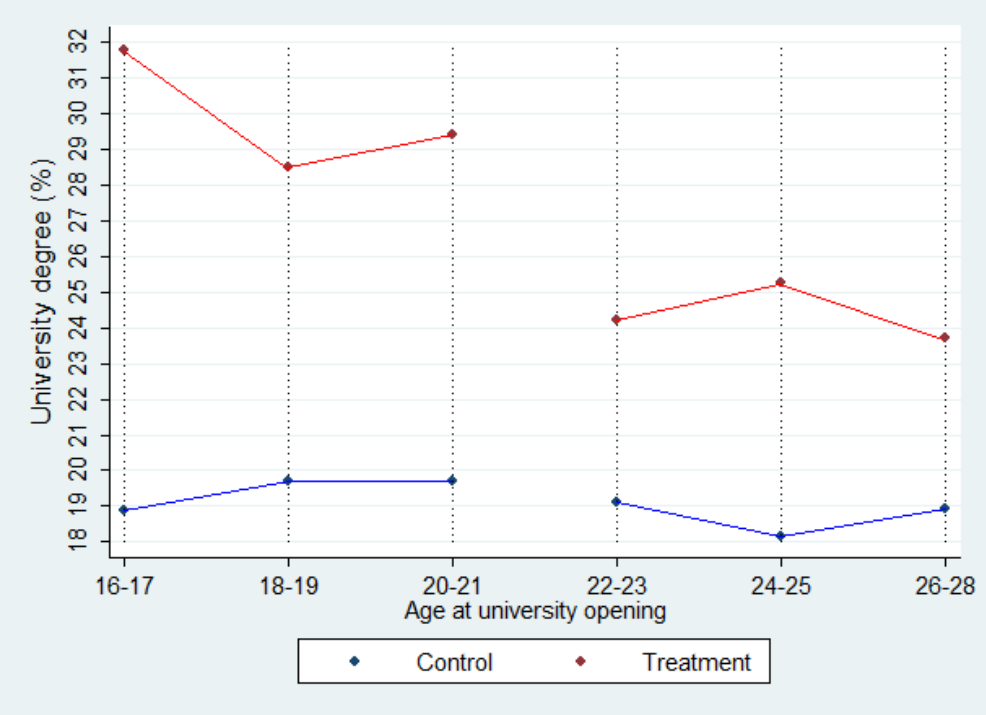
Notes: The figure shows the geographic distribution of tertiary education institutions throughout Germany. Every red flag indicates the presence of a university. Data on university openings and location from HRK (2011), own calculations.

Figure 5: Share of university graduates (green-field university openings)



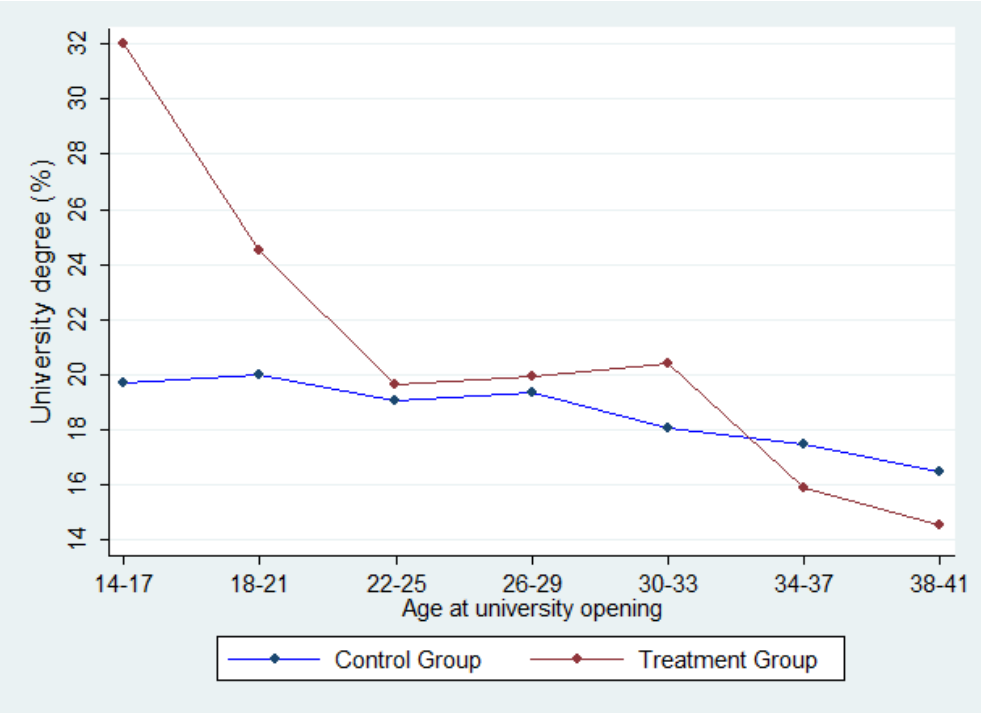
Notes: The figure shows the share of university graduates in different counties. The red line indicates individuals at the denoted age when the first university opened in their county of residence (treatment counties). The blue line indicates individuals in counties without any tertiary education institution at the same time a university opened in the treatment counties. The figure was produced using green-field university openings between 1946 and 1996. Weights were applied to account for the replication of observations. Data are from the German Socio-Economic Panel (SOEP), own calculations.

Figure 6: Share of university graduates including counties with previously existing universities



Notes: The figure shows the share of university graduates in different counties. The red line indicates individuals at the denoted age when a university opened in their county of residence (treatment counties). The blue line indicates individuals in counties with either no tertiary education institution or with at least one tertiary education institution established before 1946. The figure was produced using all university openings between 1946 and 1996. Weights were applied to account for the replication of observations. Data are from the German Socio-Economic Panel (SOEP), own calculations.

Figure 7: Trends in university completion over age groups



Notes: The figure shows the trends in the share of university graduates in different counties. The red line indicates individuals at the denoted age groups when the first university opened in their county of residence (treatment counties). The blue line indicates individuals in counties without any tertiary education institution at the time a university was opened in the treatment counties. The figure was produced using weights to account for the replication of observations. Data are from the German Socio-Economic Panel (SOEP), own calculations.

Table 1: Descriptive statistics of variables used in the analysis

Sample means by the presence of a local university

	(1) All individuals	(2) Individuals in counties with first university opening btw. 1960 - 1979	(3) Individuals in counties without any university until at least 1996	(4) Difference of means
Tertiary degree	0.202	0.239	0.199	0.041**
				(0.019)
Age (years) at university opening	21.802	21.789	21.386	0.403**
	[4.014]	[3.800]	[3.902]	(0.172)
Female	0.521	0.513	0.515	-0.002
				(0.023)
At least one parent with tertiary degree	0.102	0.085	0.075	0.010
				(0.013)
Immigrant	0.085	0.143	0.078	0.065***
				(0.016)
Observations	65201	497	36701	
Individuals	12783	497	5992	

Notes: The samples include individuals aged 16–28 at the time a new university was established. The results are weighted to account for the replication of individuals. Standard deviation in brackets. Standard error in parenthesis.

Table 2a: The effect of university opening on obtaining university degree

Dep. Var.: University degree (1=yes; 0=no)

	Green-field university openings between 1960 - 1979				
	(1)	(2)	(3)	(4)	(5)
Uni opening x Agegroup	0.100**	0.102**	0.083**	0.077**	0.080**
	(0.041)	(0.041)	(0.037)	(0.037)	(0.036)
Uni opening	-0.007	-0.008	-0.004	-0.013	
	(0.025)	(0.024)	(0.023)	(0.024)	
Agegroup	0.011	0.013*	0.012*	0.003*	0.000
	(0.007)	(0.007)	(0.007)	(0.002)	(0.001)
Female		-0.112***	-0.111***	-0.111***	-0.111***
		(0.010)	(0.010)	(0.010)	(0.009)
Parental education			0.377***	0.378***	0.343***
			(0.023)	(0.023)	(0.023)
Immigrant			-0.017	-0.022	-0.022
			(0.017)	(0.017)	(0.018)
Constant	0.193***	0.249***	0.241***	0.164***	0.042
	(0.006)	(0.009)	(0.010)	(0.025)	(0.031)
Cohort FE	No	No	No	Yes	Yes
County FE	No	No	No	No	Yes
Observations	37198	37198	37198	37198	37198
Individuals	6489	6489	6489	6489	6489
R-squared	0.003	0.022	0.093	0.099	0.177

Notes: The table shows estimation results from weighted OLS regressions. The sample includes individuals aged 16–28 at the time of a university opening. Only university openings in counties without any previous tertiary education institution are considered. Regressions in Columns (3)–(7) include a dummy variable indicating missing observations for parental education and immigrant status, respectively. Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. The total number of clusters was 3,895. * p<0.10, ** p<0.05, *** p<0.01.

Table 2b: University opening effect (different sample specification)

Dep. Var.: University degree (1=yes; 0=no)

	All openings between 1960 - 1979				
	(1)	(2)	(3)	(4)	(5)
Uni opening x Agegroup	0.056**	0.061**	0.053**	0.044*	0.053**
	(0.027)	(0.026)	(0.025)	(0.024)	(0.023)
Uni opening	0.055***	0.052***	0.042**	0.036**	
	(0.018)	(0.018)	(0.017)	(0.018)	
Agegroup	0.013*	0.014*	0.011*	0.001	-0.001
	(0.007)	(0.007)	(0.007)	(0.003)	(0.002)
Female		-0.114***	-0.115***	-0.115***	-0.115***
		(0.009)	(0.009)	(0.009)	(0.009)
Parental education			0.381***	0.381***	0.343***
			(0.020)	(0.020)	(0.020)
Immigrant			-0.019	-0.024	-0.022
			(0.015)	(0.016)	(0.017)
Constant	0.192***	0.250***	0.242***	0.243***	0.000
	(0.006)	(0.008)	(0.009)	(0.036)	(0.095)
Cohort FE	No	No	No	Yes	Yes
County FE	No	No	No	No	Yes
Observations	40815	40815	40815	40815	40815
Individuals	7630	7630	7630	7630	7630
R-squared	0.007	0.026	0.100	0.106	0.183

Notes: The table shows estimation results from weighted OLS regressions. The sample includes individuals aged 16–28 at the time of a university opening. Regressions are based on a sample considering all university openings between 1960 and 1979 (93 university openings). Regressions controlling for social background variables (parental education and immigrant status) include a dummy variable indicating missing observations for parental education and immigrant status, respectively (not shown). Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. The number of clusters was 4,452. * p<0.10, ** p<0.05, *** p<0.01.

Table 3a: University opening effect by gender

Dep. Var.: University degree (1=yes; 0=no)

	Green-field university openings between 1960 - 1979	
	Males	Females
Uni opening x Agegroup	0.029	0.140***
	(0.054)	(0.047)
Agegroup	0.002	-0.001
	(0.001)	(0.001)
Parental education	0.404***	0.287***
	(0.031)	(0.032)
Immigrant	-0.077***	0.012
	(0.029)	(0.023)
Constant	0.011	-0.273***
	(0.096)	(0.060)
Cohort FE	Yes	Yes
County FE	Yes	Yes
Observations	18194	19004
R-squared	0.216	0.188

Notes: The Table shows results for separate regressions on subsamples of males and females. The sample includes individuals aged 16–28 at the time of a university opening. Samples are based on green-field university openings between 1960 and 1979. Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. * p<0.10, ** p<0.05, *** p<0.01.

Table 3b: University opening effect by migration status

Dep. Var.: University degree (1=yes; 0=no)

	Green-field university openings between 1960 - 1979	
	Immigrants	Natives
Uni opening x Agegroup	0.179*	0.077*
	(0.094)	(0.040)
Agegroup	0.000	0.001
	(0.002)	(0.001)
Female	-0.021	-0.117***
	(0.029)	(0.010)
Parental education	0.277***	0.326***
	0.094	(0.024)
Constant	-0.096	0.047
	(0.118)	(0.033)
Cohort FE	Yes	Yes
County FE	Yes	Yes
Observations	3123	33903
R-squared	0.536	0.176

Notes: The Table shows results for separate regressions on subsamples of immigrants and natives. The sample includes individuals aged 16–28 at the time of a university opening. Samples are based on green-field university openings between 1960 and 1979. Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. * p<0.10, ** p<0.05, *** p<0.01.

Table 3c: University opening effect by parental education background

Dep. Var.: University degree (1=yes; 0=no)

	Green-field university openings between 1960 - 1979	
	High-educated parents	Low-educated parents
Uni opening x Agegroup	0.032	0.054
	(0.180)	(0.045)
Agegroup	0.000	0.000
	(0.003)	(0.001)
Female	-0.215***	-0.114***
	(0.046)	(0.012)
Immigrant	0.160	0.033
	(0.108)	(0.034)
Constant	1.215***	-0.034
	(0.046)	(0.048)
Cohort FE	Yes	Yes
County FE	Yes	Yes
Observations	2745	22063
R-squared	0.511	0.164

Notes: The Table shows results for separate regressions on subsamples of individuals with high and low educated parents. The sample includes individuals aged 16–28 at the time of a university opening. Samples are based on green-field university openings between 1960 and 1979. Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. * p<0.10, ** p<0.05, *** p<0.01.

Table 4: The effect of university opening on immobile individuals

Dep. Var.: University degree (1=yes; 0=no)

	Green-field openings between 1960 - 1979 (immobile individuals)				
	(1)	(2)	(3)	(4)	(5)
Uni opening x Agegroup	0.130**	0.130**	0.127**	0.129**	0.146***
	(0.065)	(0.064)	(0.060)	(0.059)	(0.052)
Uni opening	0.028	0.026	0.013	-0.001	
	(0.037)	(0.036)	(0.036)	(0.035)	
Agegroup	0.006	0.006	0.004	0.003	0.000
	(0.009)	(0.009)	(0.009)	(0.003)	(0.001)
Female		-0.073***	-0.074***	-0.076***	-0.068***
		(0.015)	(0.014)	(0.014)	(0.014)
Parental education			0.354***	0.350***	0.302***
			(0.047)	(0.047)	(0.047)
Immigrant			0.252*	0.267*	0.229*
			(0.141)	(0.139)	(0.125)
Constant	0.116***	0.154***	0.145***	0.182**	-0.059
	(0.009)	(0.012)	(0.013)	(0.075)	(0.099)
Cohort FE	No	No	No	Yes	Yes
County FE	No	No	No	No	Yes
Observations	12265	12265	12265	12265	12265
Individuals	2140	2140	2140	2140	2140
R-squared	0.009	0.021	0.083	0.100	0.259

Notes: The table shows estimation results from weighted OLS regressions. The sample includes individuals aged 16–28 at the time of a university opening who indicated still living at the place of their childhood. Only university openings in counties without any previous tertiary education institution are considered. Regressions in Columns (3)–(7) include a dummy variable indicating missing observations for parental education and immigrant status, respectively. Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. The total number of clusters was 1,942. * p<0.10, ** p<0.05, *** p<0.01.

Table 5: The effect of university opening using a broader sample

Dep. Var.: University degree (1=yes; 0=no)

	Green-field openings between 1946 - 1996					All openings between 1946 - 1996				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Uni opening x Agegroup	0.057*	0.058*	0.050	0.044	0.040	0.052**	0.054**	0.045**	0.039*	0.038**
	(0.032)	(0.032)	(0.031)	(0.031)	(0.029)	(0.022)	(0.022)	(0.021)	(0.020)	(0.019)
Uni opening	0.010	0.010	0.002	-0.007		0.055***	0.054***	0.040***	0.030**	
	(0.021)	(0.021)	(0.020)	(0.020)		(0.015)	(0.015)	(0.014)	(0.014)	
Agegroup	0.003	0.004	-0.001	0.003	0.003*	0.006	0.006	0.002	0.001	0.002
	(0.006)	(0.006)	(0.005)	(0.002)	(0.001)	(0.006)	(0.006)	(0.005)	(0.002)	(0.002)
Female		-0.084***	-0.082***	-0.080***	-0.080***		-0.087***	-0.085***	-0.084***	-0.085***
		(0.007)	(0.007)	(0.007)	(0.007)		(0.007)	(0.007)	(0.007)	(0.007)
Parental_Education			0.340***	0.344***	0.318***			0.347***	0.350***	0.321***
			(0.016)	(0.016)	(0.016)			(0.014)	(0.014)	(0.014)
Immigrant			-0.015	-0.017	-0.024*			-0.019	-0.022*	-0.026**
			(0.012)	(0.012)	(0.013)			(0.012)	(0.012)	(0.012)
Constant	0.189***	0.232***	0.213***	0.329*	0.045	0.188***	0.233***	0.212***	0.282*	-0.024
	(0.005)	(0.007)	(0.007)	(0.194)	(0.124)	(0.005)	(0.006)	(0.007)	(0.161)	(0.108)
Cohort FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
County FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	59458	59458	59458	59458	59458	65201	65201	65201	65201	65201
Individuals	11015	11015	11015	11015	11015	12783	12783	12783	12783	12783
R-squared	0.001	0.012	0.081	0.090	0.153	0.005	0.017	0.090	0.098	0.161

Notes: The table shows estimation results from weighted OLS regressions. The samples include individuals aged 16–28 at the time of a university opening. Regressions shown in Columns (1)–(5) are based on a sample considering only university openings between 1946 and 1996 that were the first tertiary education institution in a county. In Columns (6)–(10), all university openings between 1946 and 1996 are considered. Regressions controlling for social background variables (parental education and immigrant status) include a dummy variable indicating missing observations for parental education and immigrant status, respectively (not shown). Standard errors (in parentheses) were clustered to account for dependence of observations within clusters of county and year of birth. The number of clusters was 6,790 in regressions (1)–(5), and 7,681 in regressions (6)–(10). * p<0.10, ** p<0.05, *** p<0.01.

Table A1: Descriptive statistics for subsample of immobile individuals

Sample means by the presence of a local university

	(1) Immobile individuals in counties with first university opening btw. 1960 - 1979	(2) Immobile individuals in counties without any university until at least 1996	(3) Difference of means
Tertiary degree	0.200	0.133	0.067**
			(0.030)
Age (years) at university opening	21.706	21.308	0.397
	[3.747]	[3.912]	(0.281)
Female	0.461	0.504	-0.043
			(0.037)
At least one parent with tertiary degree	0.067	0.062	0.005
			(0.019)
Immigrant	0.017	0.006	0.011
			(0.010)
Observations	180	14310	
Individuals	180	2381	

Notes: The samples include individuals aged 16–28 at the time a new university was established. The results are weighted to account for the replication of individuals. Standard deviation in brackets. Standard error in parenthesis.