working paper

Österreichisches Institut für Familienforschung
Austrian Institute for Family Studies

Nummer **27 – 2002**

Titel **"UNIVERSITY EDUCATION – AN ANALYSIS IN THE CONTEXT OF THE FAMSIM+ MICROSIMULATION MODEL FOR AUSTRIA"**

Autoren:
Franz Schwarz
Martin Spielauer
Karin Städtner

working papers have only received limited review

ÖIF, Gonzagagasse 19/8, A-1010 Wien
Tel. +43-1-535 14 54-0
Fax +43-1-535 14 55
url: http://www.oif.ac.at
email: team@oif.ac.at

Content

Abstract

This working paper on university education analyzes the influencing factors on university enrollment rates, graduation vs. dropout rates as well as study durations in Austria. The analysis is based on the special program of the 1996 micro census that includes detailed educational histories. On the macro level, a considerable increase of enrollment rates can be observed that more than doubled in the second half of the last century. Female enrollment rates that were only one third of the male rates for the 1935-39 birth cohort, have already drawn level with the male rates. Changes on the macro level are the outcome of decisions taken by the individual agents on the micro level. Moving from the macro to a micro analysis of university enrollments and graduations, a very stable relation between parents' educational attainments and the university graduation rates of their offspring can be found, besides the diminishing gender differences in educational behavior. For given educational levels of the parents, enrollment rates almost remained constant or even slightly decreased over time for the males, while female rates moderately increased and eventually met the male rates in the last decades. Given the marked differences in enrollment rates according to parents' education – for the last birth cohort 1965-69 the rates range from around 4% to nearly 60% for the lowest and the highest (of the five) educational groups – a considerable part of the changes on the macro level can therefore be attributed to the changing educational composition of the parents' generation. This result is consistent with the findings regarding prior educational transitions as studied in two previous papers. Being written in the framework of the FAMSIM+ microsimulation model currently under development at the Austrian Institute for Family Studies, various regression models regarding university education will be developed in this paper, besides providing a descriptive analysis. These models will serve as basis of the behavioral modules of the FAMSIM+ model that consists of three sub modules regarding university enrollments and graduations: First, it is determined whether a person enrolls at university. Second, it is decided if an enrolled person will graduate from university or drop out. Third, the individual enrollment duration is determined by means of a survival analysis.

1 Introduction

This working paper on University Education is the third paper analyzing educational careers in the context of the FAMSIM+ microsimulation model for Austria. In our preceding papers we have carried out descriptive analyses and developed logistic regression models at the transition from primary to lower secondary education (Spielauer et. al., 2002) and from lower to upper secondary education (Schwarz, et. al., 2002). This paper concludes the “trilogy” of papers analyzing the career paths of individuals within the Austrian School system.

As in the previous papers this analysis is based on the special program of the 1996 micro census that includes detailed educational histories. In contrast to primary and secondary education, which is usually completed between the ages of 6 and 18, university studies are not completed within a predetermined period or started at a specific age. As the foregoing chapters have shown, the duration for completing university studies varies from 3 to 10 years or more, and the enrollment age differs strongly as well. Additionally, some individuals decide to work after finishing upper secondary education and choose to enroll at university at a later stage. In such a case the admission qualification for university enrollment is often reached via second-chance education, which can be a *Matura* evening school, a 'Berufsreifeprüfung' or a 'Studienberechtigungsprüfung'. Because of the effect on the pattern of life, such as childbearing, we are interested in modeling first-chance education. Thus, individuals who studied in second-chance education are not of interest in our analysis; nevertheless, the modeling of second-chance education will be considered at a later date. For that reason, we restricted the university enrollment age to 24 years. This allows a longer duration in upper secondary education if, for instance, individuals attend several types of upper secondary education, as well as the fact that males have to fulfill compulsory military service.

Since not only the background plays an essential role, for instance the education of parents, but the prior life course as well, in the Family Microsimulation Model the previous educational path is taken into account too. In general, there are two ways to enter university education: from secondary vocational and technical colleges (BHS) and from secondary academic schools (AHS), both completed with a *Matura* degree. The reason to separate *Matura* into the two different school types lies in the completely different objectives of these schools. While secondary vocational and technical colleges provide individuals with vocational knowledge, secondary academic schools prepare students especially for university education. Hence, for academic school graduates we can expect considerably higher probabilities for university enrollments. Consequently, logistic regression analysis delivers for both admission options a regression equation, which enables us to calculate the transition probabilities from AHS, or BHS respectively, to university for any combination of the influence factors *education of parents*, *residential area* and *gender*.

Since the university education module of the FAMSIM+ Project consists of several steps, we first estimate the probability of university enrollment by means of logistic regression. By using the Monte Carlo simulation, we are able to decide whether or not individuals enter university education. Second, the risk of dropping out is estimated by logistic regression. Third, we have to determine the duration of the educational process until graduation or dropout respec-

tively. This is done with a survival analysis, where we perform for both outcomes, graduation and dropout, two separate analyses. By doing so, we obtain for every point in time the probability of dropping out or graduating respectively.

However, before modeling university education by logistic regression and survival analysis, a descriptive look into the matter is carried out. First, influence factors on university enrollment and graduation and changes over time are investigated, followed by a look into the factors that affect dropout rates. Afterwards, enrollment and graduation age is examined. In the last chapter of the descriptive analysis durations until graduation are studied.

2 Data and Variables

The data source for this study is the special program of the Austrian micro census from June 1996, which contained a questionnaire on educational history, marriage, and biography of births. For our analysis the questions of interest were:

- Kind of graduation in compulsory education
- All kinds of ever attended/started educations following compulsory education
- Educational history starting from lower secondary education
- Highest education of the individual's parents
- Municipality type: rural or urban (place where person lived at the age of fifteen)

In contrast to the basic program of the micro census, the special program is voluntary. Since for various reasons individuals refuse to answer the questions, e.g. due to embarrassment or lack of interest, we have to consider a systematic error. Particularly in a survey on education, we can expect less interest in the program from individuals with a lower educational level; consequently, some caution in usage of the results is advisable. In our sample 26 639 out of 33 811 individuals answered the questions on education. However, since we evaluate individual behavior in dependency on influence factors, it is not so important that a certain proportion in the sample, i.e. the proportion of university enrollments, is equivalent to the proportion in the total population, but that the individual behavior in the sample corresponds to the behavior of the individuals in the total population. For instance, the sample proportion of university enrollments of individuals of parents holding a university degree should be approximately equivalent to the respective proportion in the total population. Especially for the logistic regression and the survival analysis, carried out in this paper, the systematic error is basically irrelevant, as the non-respondents display the same behavior as the respondents concerning an obligatory basic program and a voluntary special program.

Table 2.1: Variable description

Variable	Description	Type
birth_dat	Date of birth	Numerical
gender	Gender	Categorical with the categories 0 female 1 male
municip	Municipality type, where lived at age of fifteen	Categorical with the categories 0 rural 1 urban
educ_com	Compulsory education	Categorical with the categories 0 Lower secondary school ('Haupt-schule') 1 Lower academic secondary school ('AHS')
educ_hgh	Highest education of the individuals inter-viewed	Categorical with the categories 1 compulsory 2 apprenticeship 3 vocational 4 *Matura*[1] 5 university
educ_par	Highest education of parents (highest value of father or mother)	c.f. *educ_hgh*
uni_educ	University graduate according the basic program of the micro census	Categorical with the categories 1 yes 0 no
univ_enr	Ever enrolled at university	Categorical with the categories 1 yes 0 no
univ_adm	Admission qualification (previous secondary education)	Categorical with the categories 4 secondary academic school (AHS) 5 vocational and technical school (BHS)
uni_beg	Year of university enrollment	Numerical
uni_end	Year of leaving from University	Numerical
uni_dur	Duration till university leaving (gradation or drop out) in years	Numerical
is_drop	Kind of university leaving	Categorical with the categories 1 dropout 0 graduation
gew1	Adjusted weight by the population structure of the total population of Austria	Real

[1] In Austria, the '*Matura'* (or '*Reifeprüfung'*) is the final exam in upper level academic secondary schools and secondary technical and vocational colleges, usually taken after 12, or respectively,. 13 years of education. *Matura* also grants admission to universities or Fachhochschulen (post-secondary colleges).

3 Influence Factors on University Education

The main purpose of this chapter is to answer the question of who studied, graduated or dropped out in relation to the influence factors gender, municipality type, compulsory education, secondary education and education of parents, looking at all individuals interviewed. Table 3.1 gives a concise overview of sample sizes and distributions of university studies in dependency on these influence factors for individuals born between 1937 and 1981. From the 13 220 females in our sample, 3.7% already graduated, 1.4% dropped out and 2.3% are still enrolled; hence, 7.4% respondents answered that they were/are enrolled while the rest of 92.6% was/is not. Comparing the statements of females with those of their male counterparts, Table 3.1 reveals great gender differences. While 5.9% of males responded that they were university graduates, only 3.7% of females did so.

Regarding residential area, clear differences between rural and urban areas emerge. Only 5.2% of the individuals living in rural areas ever enrolled at university, of whom 2.6% graduated. In contrast, the percentages for individuals in urban areas are approximately three times higher.

As expected after the results of the two foregoing working papers concerning Austrian education (Spielauer et. al., 2002; Schwarz et. al., 2002), the previous educational path influences the probability of attending university to a great extend. Adding graduates, dropouts and currently enrolled, only 1.7% of all individuals who attended lower secondary school (Hauptschule) enrolled at university, whereas 35.7% did so if they attended lower secondary academic school (AHS).

In order to get admission to university a *Matura* certificate is required. This is normally gained by successfully graduating from vocational and technical college or from secondary academic school. However, those who did not obtain the necessary *Matura* certificate within secondary education but who nevertheless want to attend university, can acquire their admission qualification via second-chance education. For the purpose of this working paper we neglected the second possibility and only concentrated on the 'regular' admission to university. When studying the admission qualification shown in Table 3.1, clear differences between vocational and technical college graduates and secondary academic school graduates became visible. While 57.4% of secondary academic school graduates stated that they were enrolled at university, only 26.9% of vocational and technical college graduates did so.

As became also evident in the two foregoing working papers, the education of parents considerably influences the educational path of their children. These findings are also relevant at university level. While the vast majority (97.3%) of individuals whose parents have compulsory education only as their highest educational attainment never enrolled, this proportion decreased gradually with the education of parents.

Table 3.1: University studies in dependency on influence factors for individuals born between 1937 and 1981

Variables	Categories	n	never enrolled	university graduates	university dropouts	currently enrolled
			in %			
Gender	Female	13220	92.6	3.7	1.4	2.3
	Male	13419	89.7	5.9	1.4	3.1
Municipality type	Rural	17551	94.8	2.6	0.7	1.8
	Urban	7786	85.0	8.6	2.5	3.9
Compulsory education	Lower secondary school	20637	98.3	0.8	0.3	0.6
	Lower secondary academic school	4384	64.3	19.3	5.6	10.8
Admission qualification	Vocational college (BHS)	2045	77.1	9.0	4.3	9.6
	Academic school (AHS)	2647	42.6	33.8	8.1	15.5
Education of parents	Compulsory	11708	97.3	1.8	0.4	0.5
	Apprenticeship	7034	94.4	2.7	1.1	1.8
	Vocational school	1818	85.2	6.9	2.9	5.0
	Matura	1826	73.1	15.0	4.2	7.7
	University	1272	52.9	26.2	5.3	15.6

4 Changes over Time for University Enrollments and Graduations

At this point we consider changes over time in relation to several influence factors. In order to prevent too small a sample size in the cohorts and censured data of individuals who are still in secondary education, we restricted the time span to 1935-1969, creating five-year birth cohorts for this time period. Nevertheless, in the last cohort we still have to consider a slight data censoring problem.

4.1 Changes over Time Concerning University Enrollments

In this chapter enrollments – independent of the fact of how long the individuals were enrolled and whether or not they were successful – will be studied closely. By comparing birth cohorts structural changes and time trends shall be discovered.

When looking at Figure 4.1 it becomes evident that the total number of individuals enrolled at university increases enormously over the four birth cohorts considered, with a particularly strong increase from the 50s to the 60s. While only 5% of the individuals of the cohort 1935 to 1939 ever enrolled, this percentage tripled to 15% for the final cohort observed.

Figure 4.1: Differences between birth cohorts in the proportion of individuals enrolled at university in dependency on gender

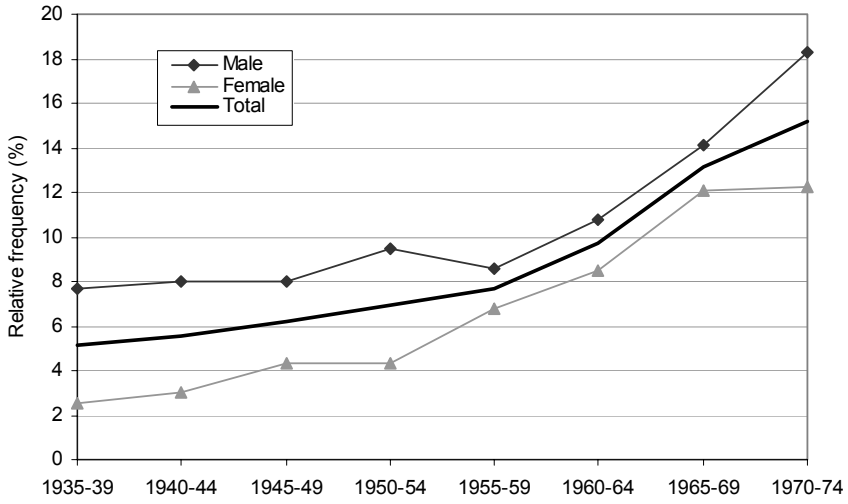
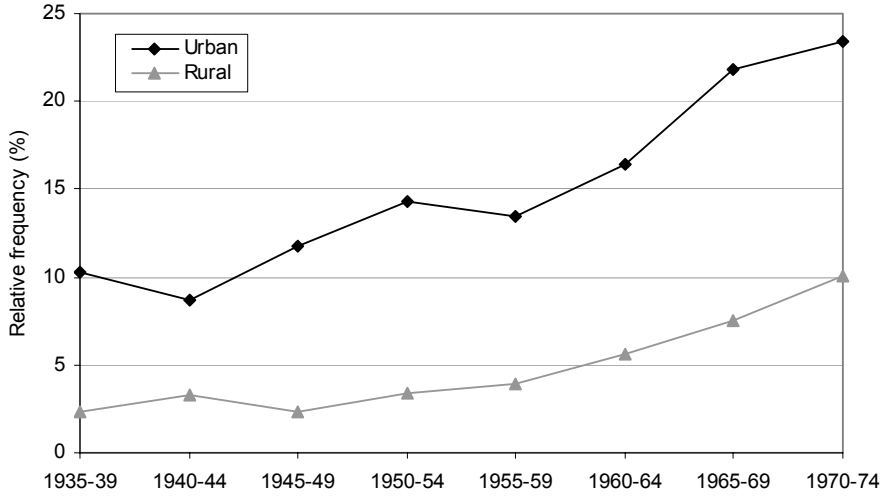


Figure 4.1 also shows gender differences and their development over time. Although the development of the two graphs is quite similar, the graph concerning females shows a more marked increase. Therefore the initial gender gap of about five percentage points narrowed to two percentage points for the birth cohort 1965-69. The gap is widening in the last cohort, which is probably due to the censored data.

Figure 4.2 makes clear that there is a substantial difference between the residential areas where individual lived at the age of 15 and the enrollment rate at university. Although in both the rural and urban areas there has been a steady rise in the number of individuals enrolled, the increase was far more pronounced in urban areas which made the graphs diverge fur-

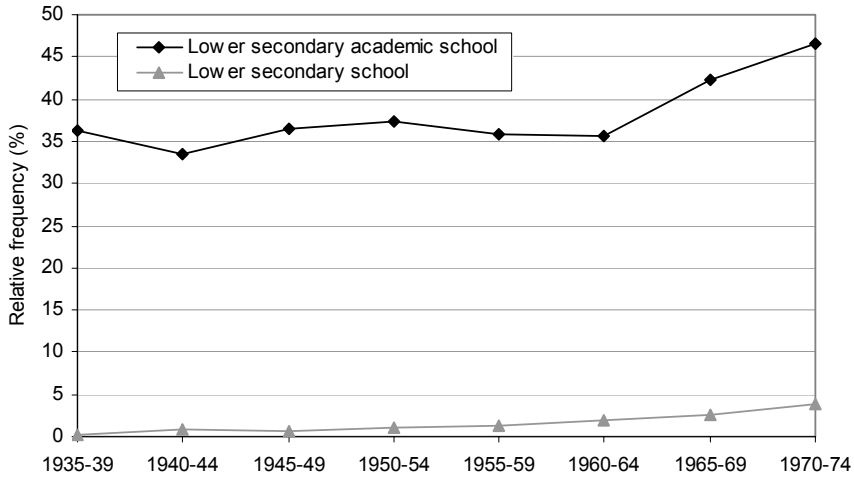
ther. For the final cohort the corresponding relative frequencies are 23% for individuals living in urban regions and 10% for those in rural areas.

Figure 4.2: Differences between birth cohorts in the proportion of individuals enrolled at university in dependency on municipality area



As the previous paper on upper secondary education has shown, lower secondary education influences the educational future of individuals to a great extent. Figure 4.3 displays that the probability of university enrollment for individuals with 'Hauptschule' is small; however, for these individuals an extensive rise from 1% to nearly 8% took place over the last decades. Nevertheless, the gap between these two school types is still substantial, since for lower secondary school (AHS) the enrollment rate was already 35% in the past and increased further to around 45%.

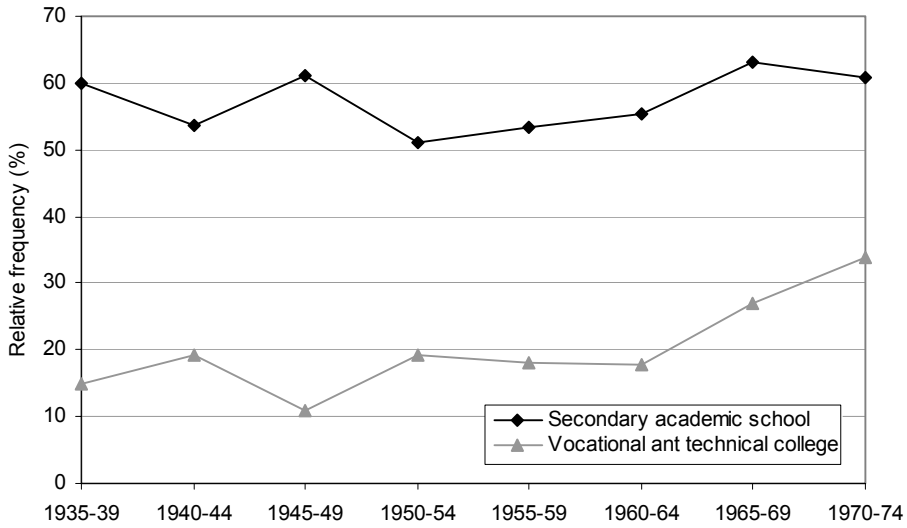
Figure 4.3: Differences between birth cohorts in the proportion of individuals enrolled at university in dependency on lower secondary education



As emerges from Figure 4.4, time trends regarding the relative frequency of becoming a student differ depending on the type of secondary education completed. There appears a stagnation at around 60% for individuals enrolled at university who previously attended secondary academic school. On the other side, there was a constant rise in the proportion of

graduates from vocational and technical colleges enrolled at university, which more than doubled from around 15% to over 30%.

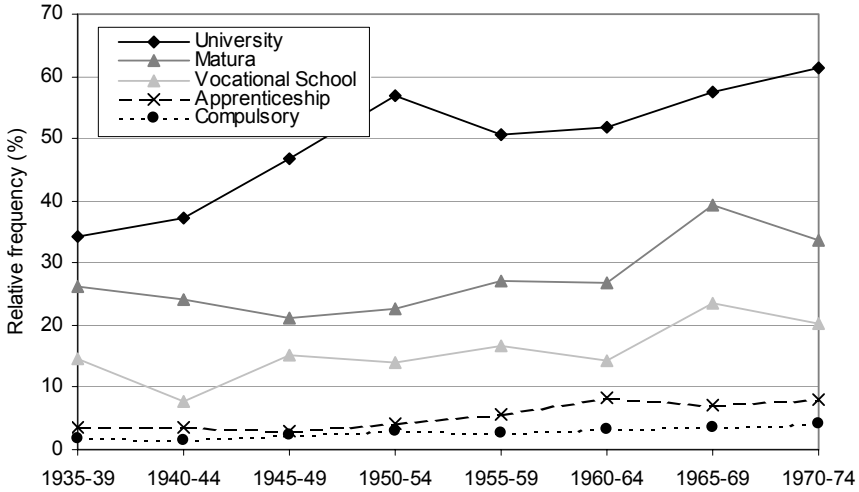
Figure 4.4: Differences between birth cohorts in the proportion of individuals enrolled at university in dependency on secondary education



As we found in our previous research, the educational attainment of parents plays a crucial role when investigating education. Therefore, we want to examine whether this applies to the university level as well. Figure 4.5 shows variations over time regarding parental schooling and the proportions of individuals enrolled at university. After an initial downward movement, a general upward trend can be observed for all groups defined in our study.

If parents are university graduates, individuals often also get/take the chance to study. When comparing individuals whose parents are university graduates and who are born within the period of 1940-44 with those born between 1965-69, a rise in the relative frequency of enrollment from 37% to 58% can be observed. However, the number of students of parents who have compulsory education only almost doubled and reached a maximum value of 4% for the final cohort. Thus, over time the disparities regarding the relation between enrollments and parental education experienced a reinforcement.

Figure 4.5: Differences between birth cohorts in the proportion of individuals enrolled at university in dependency on the education of parents



4.2 Changes over Time Concerning University Graduates

As so far we have only been interested in enrollments, regardless of whether or not the individuals enrolled actually graduated from university, the central aim of the following considerations is to answer the question of whether the variables of interest influenced the proportion of graduates over time. In the following analysis it has to be kept in mind that some individuals of the birth cohorts 1960-64 and 1965-69 might still have been studying in 1996, when the micro census was conducted. Therefore, our results may be biased downward, which we nevertheless accept, as we do not want to omit these observations. However, with these limitations in mind, the subsequent figures have to be interpreted cautiously.

As Figure 4.6 reveals, there has been a steady increase in the overall relative frequency of university graduates, which is mainly due to female enrollment rates. The gender gap of university graduates was reduced from almost 4.5 percentage points for the birth cohort 1930-1934 to 1.5 percentage points for the cohort 1955-59. However, the percentage for the years of birth 1960-69 may be censured, as not all individuals born during this time-interval may already have graduated.

Figure 4.6: Differences between birth cohorts in the proportion of individuals graduated from university in dependency on gender

Figure 4.7 shows the considerable differences between urban and rural areas concerning university graduates. While a moderate rise from around 8% to about 10% in the proportion of graduates from urban areas took place, the proportion of graduates coming from rural areas doubled from around 2% to 4%.

Figure 4.7: Differences between birth cohorts in the proportion of individuals graduated from university in dependency on municipality area

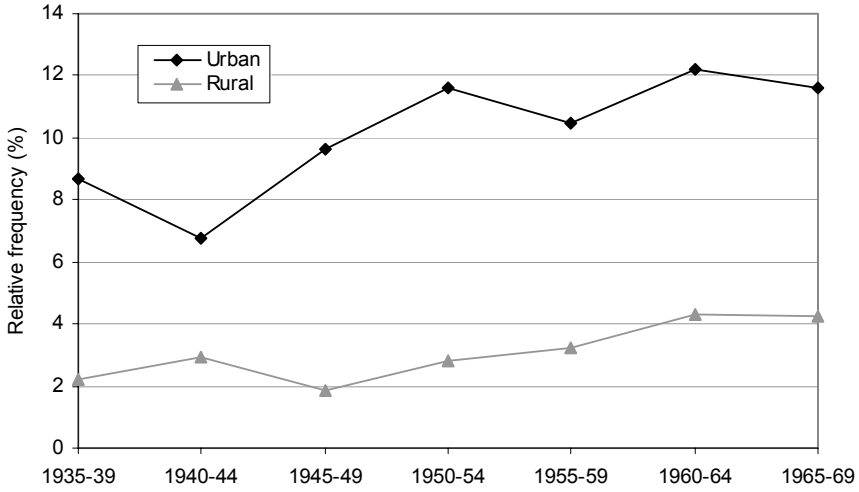


Figure 4.8 shows the proportion of graduates and their development over birth-cohorts with regard to their compulsory education. While the relative frequency of lower secondary school graduates becoming university graduates increased only slightly to a level of 1% for the last cohort, the graph applying to lower secondary academic school graduates dropped from around 30% to 25%. In the last cohorts this downward movement may be traced back to the fact that not all individuals had finished their studies at the time the data was collected.

Figure 4.8: Differences between birth cohorts in the proportion of individuals graduated from university in dependency on lower secondary education

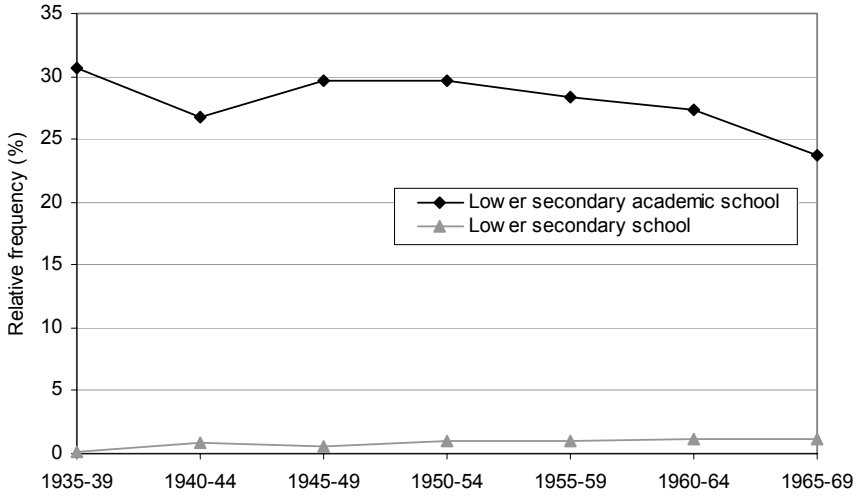
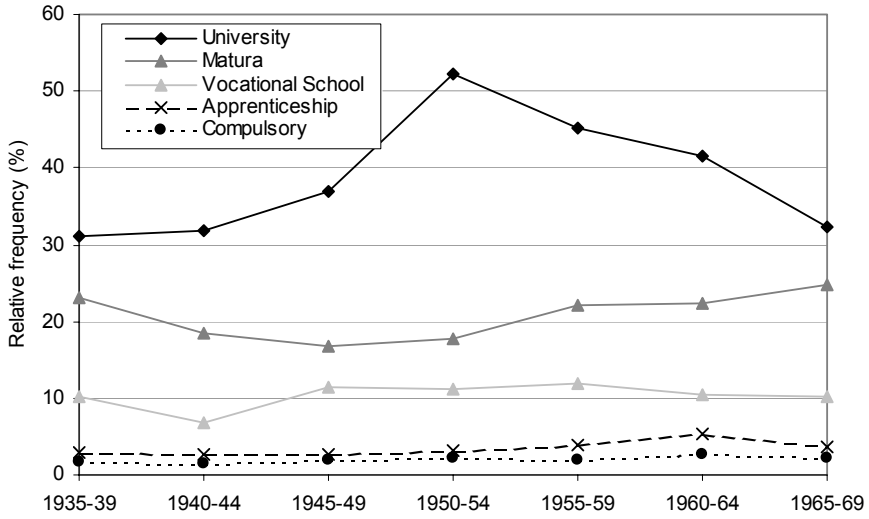


Figure 4.9 pictures a steadily declining proportion of graduates concerning individuals graduating from secondary academic schools. The graph for those individuals who graduated from vocational and technical colleges remains nearly constant.

Figure 4.9: Differences between birth cohorts in the proportion of individuals graduated from university in dependency on secondary education

Figure 4.10 shows graduation rates in relation to the education of parents. For individuals whose parents have a *Matura* degree or attended a vocational school, no clear time trends can be identified. A substantial increase in the proportion of graduates took place for those individuals having parents who hold a university degree, which peaked in the cohort 1950-54. For individuals with parents who have compulsory education or did an apprenticeship, a negligible increase in the proportion of graduates took place over the cohorts investigated, whereas the ratio of graduates whose parents attended vocational school remained more or less constant at 10% over the entire period observed.

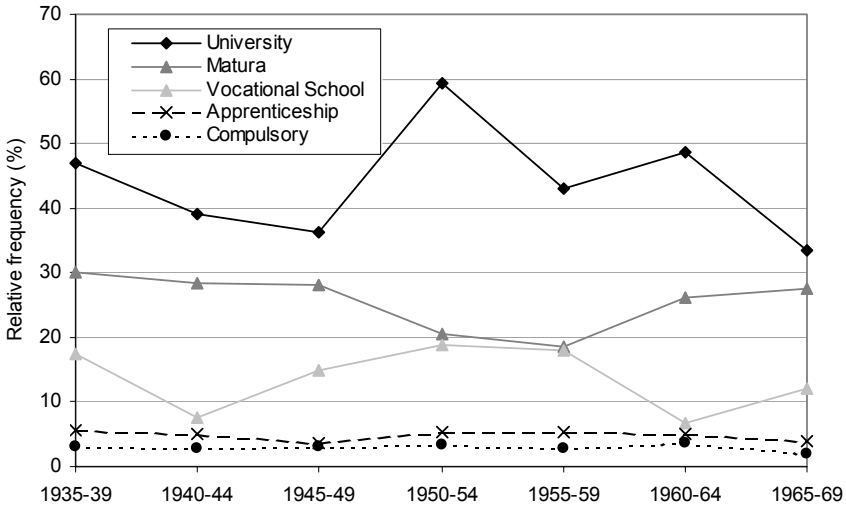
Figure 4.10: Differences between birth cohorts in the proportion of individuals graduated from university in dependency on educational attainment of parents



The considerable differences in enrollment rates according to parents' education – for the last birth cohort 1965-69 the rates range from around 4% to nearly 60% for the lowest and highest (of the five) educational groups – indicate that a considerable part of the overall increase of graduation rates observed on the macro level can be attributed to the changing educational composition of the parents' generation.

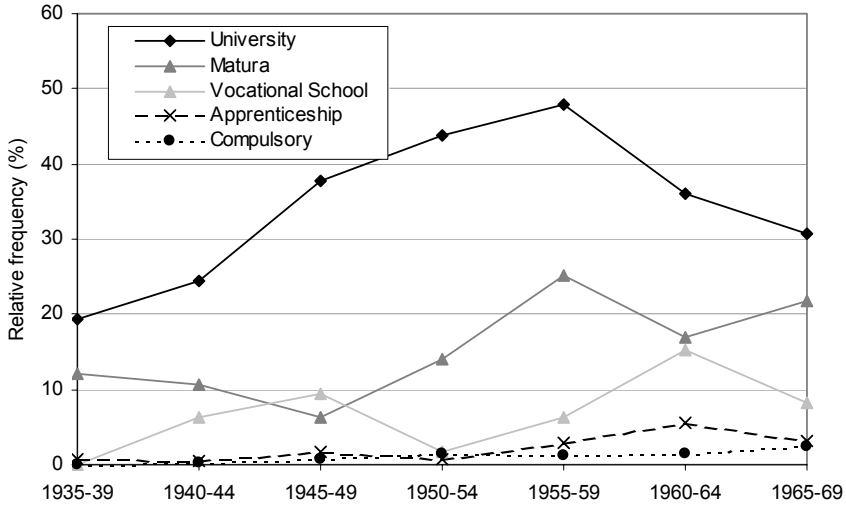
A rather stable behavioral relation between parents' education and graduation rates of offspring can be found for males. As shown in Figure 4.11, graduation rates by parents' education remained almost stable over all decades for parents having compulsory education or apprenticeship. For all others no clear time trends can be observed.

Figure 4.11: Differences between birth cohorts in the proportion of MALE individuals graduated from university in dependency on educational attainment of parents



However, this picture changes notably for the female population. As shown in Figure 4.12, female graduation rates increased considerably for all groups and drew level with the male rates. However, it has to be mentioned that the sample size is rather small for female graduates devided by the education of parents.

Figure 4.12: Differences between birth cohorts in the proportion of FEMALE individuals graduated from university in dependency on educational attainment of parents



5 University Dropout Rates

In this chapter we examine university dropout rates in relation to the proportion of graduates, which is, of course, the contra probability of the dropout rates ($p_g = 1 - p_d$). First, we will consider the several influence factors on university graduation for individuals who left university between 1966 and 1995. Second, we will look at changes in the dropout rates over the last decades. Since often individuals do not state uncompleted studies, differences between our rates and those of university statistics can occur. Additionally, we did not count individuals as dropouts if they enrolled again within two years of their first interruption and graduated in the second attempt.

Table 5.1: Influence factors on university leaving for individuals who left university between 1966-1995

Variable	Category	n	Graduates	Dropouts
			in %	
Gender	Female	467	74.1	25,9
	Male	731	81.3	18,7
Municipality type	Rural	436	77.7	22,3
	Urban	705	77.7	22,3
Compulsory education	Lower secondary school	189	70.8	29,2
	Lower secondary academic school	887	78.1	21,9
Admission qualification	Secondary academic school	914	81.2	18,8
	Vocational and technical college	229	67.0	33,0
Education of parents	Compulsory	190	81.7	18,3
	Apprenticeship	210	70.4	29,6
	Vocational school	125	70.8	29,2
	Matura	271	79.2	20,8
	University	308	83.6	16,4
Total		1198	78,4	21.6

As in the preceding chapters, we will determine the extent by which gender, municipality type, compulsory school, secondary school and the education of parents influence the probability of becoming a dropout, or graduate respectively. Table 5.1 reveals that there are gender differences regarding the two ways of leaving university. While 18.7% of male students drop out of university education, 25.9% of female students do so. No distinction can be drawn between rural and urban municipalities where both show a dropout rate of 22.2%. Concerning compulsory education, the relative frequency of dropping out of university is 7.3 percentage points higher for students previously attending lower secondary school than for those attending lower secondary academic school.

Nevertheless, notable differences can be observed as regards the admission qualification, which describes the way by which the individual acquired the qualification necessary for attending university. Among vocational and technical school graduates the dropout rate is 33% and, thus, quite high compared to students who gained their admission qualification by graduating from secondary academic school (18.8%).

An interesting correlation arises regarding the educational attainment of parents and the percentages of graduates or dropouts respectively. While the dropout rate is quite similar and lies at around 30%, both for parents previously doing an apprenticeship or attending a vocational school, it is ten percentage points smaller for individuals whose parents attended compulsory school only or hold a *Matura* degree. As expected, with 16.4% the far lowest dropout rate applies to those individuals whose parents hold a university degree.

Looking into changes over time, Figure 5.1 shows the development of dropouts for females and males over the past 35 years. In general, the rates increased during the last decades. The unstable behavior of the graphs may partly be traced back to the fact that we used event cohorts of dropping out or graduating from university respectively in order to avoid censured data. In such a case, the dominator of the rates may be biased to some extent, if in (a) specific year(s) the number of enrollments was much higher or lower, since the event 'dropout' occurs some years earlier than the event 'graduation'. However, as we examined birth cohorts too, the unstable behavior was reduced markedly. Additionally, we have to consider the small sample size.

The peak value for females leaving university between 1986-90 is noticeable in Figure 5.1. This high rate deserves a closer look which is taken in Figure 5.2. We are able to confine the peak to individuals having BHS as admission qualification. If we subdivide the category BHS by gender, the high values can be traced back to females with BHS; although we have to consider the small sample size at this point. In contrast to BHS, the rates of individuals from AHS show a stable behavior with negligible changes over time.

Figure 5.1: Dropout rates subject to gender over 5-year cohorts of university leaving

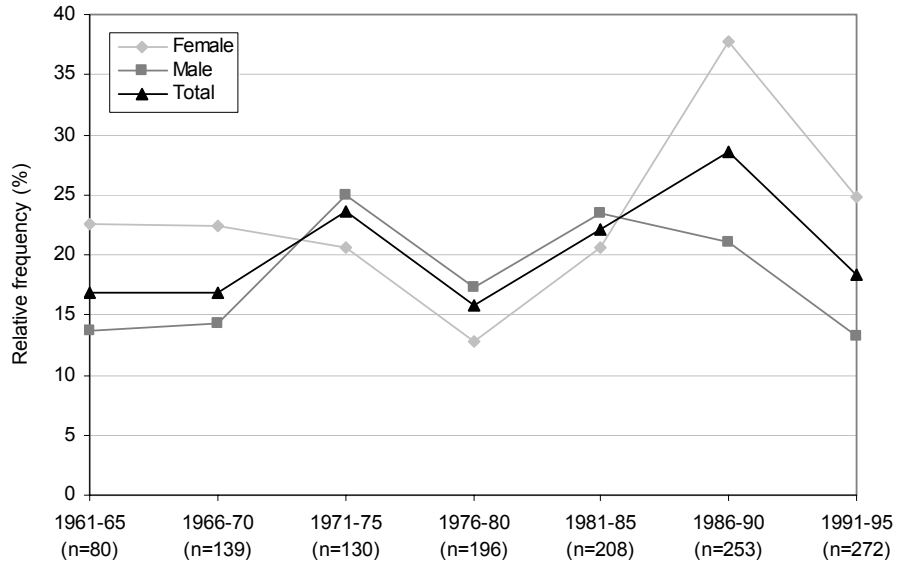
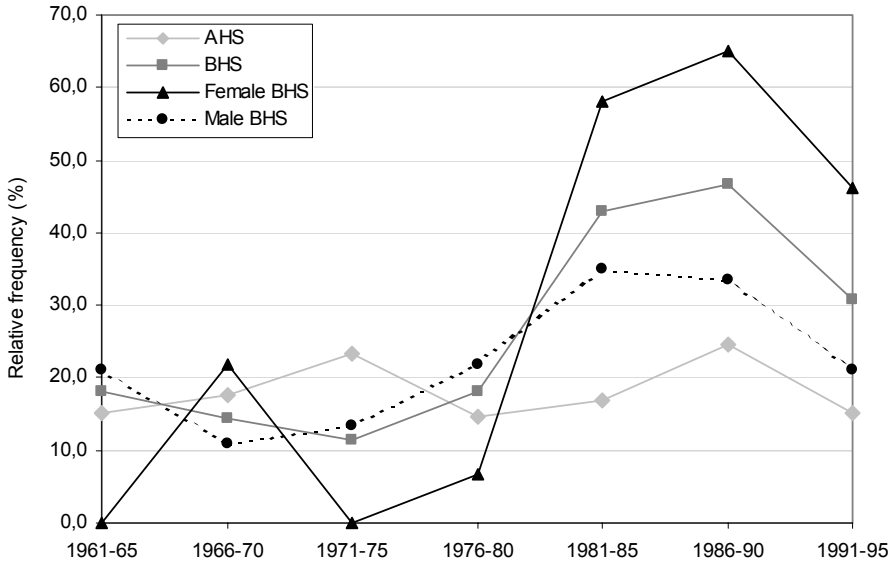
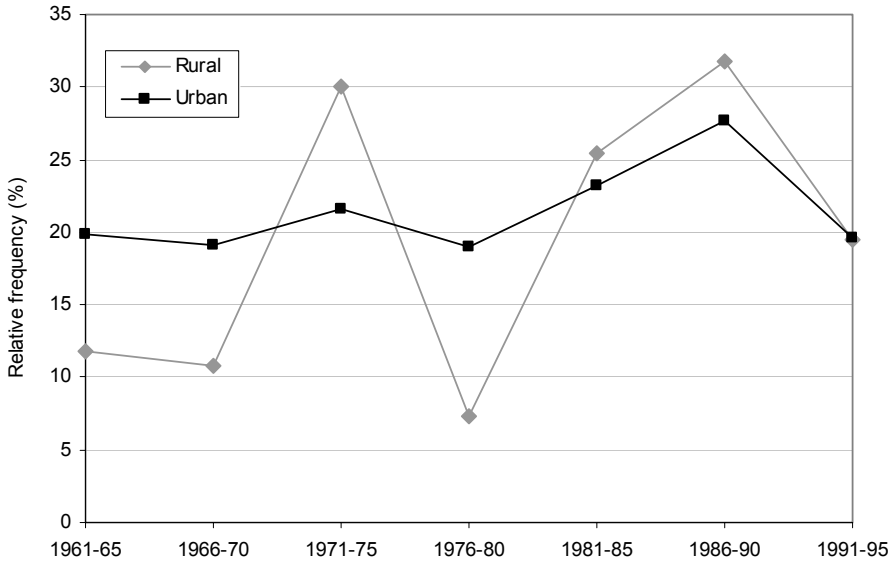


Figure 5.2: Dropout rates subject to admission qualification in relation to gender and BHS over 5-year cohorts of university leaving



For residential area an unstable behavior of the dropout rates applies to rural municipalities; however, if we smooth the graphs, we observe similar dropout rates that increase slightly for both areas, rural and urban.

Figure 5.3: Dropout rates subject to residential area over 5-year cohorts of university leaving



6 Enrollment and Graduation Age

While birth cohorts and the population as a whole were the matter of interest of the previous sections, we now concentrate on individuals enrolled as well as on graduates. Thus, in this chapter we no longer regard birth cohorts, but 'enrollment' and 'graduation cohorts'. These two cohorts consist of all individuals who enrolled or respectively graduated during the specified period, independent of the fact when they were born. Thus, enrollment and graduation cohorts are not entirely comparable, as they partly consist of different individuals. When comparing different enrollment and graduation cohorts, one has to be aware that the sample size reduces considerably, the further we go back in time.

In the following analysis the age at the time of enrollment or respectively graduation applies to first studies only. We also investigated all individuals who enrolled and dropped out, if they enrolled again within two years of their interruption. Thus, we allow for students that changed their studies, which is fairly common in Austria. Because we are only interested in modeling first-chance education, we only consider individuals who enrolled between the ages of 17 and 24. These adjustments were necessary, as the inclusion of higher enrollment age and, thus, second-chance education would distort our results.

6.1 Changes Over Time in Enrollment and Graduation Age

As Table 6.1 reveals that the median age at the time of enrollment hardly changed during the 40-year-period considered. It remained more or less constant at a value of about 228 months – equivalent to 19 years – for the four graduation cohorts. This indicates that most individuals started their university career immediately after completing secondary education.

Table 6.1: Age at enrollment in years over 10-year enrollment cohorts from 1956 to 1995

Enrollment cohorts	n	Median	Interquartile range[2]	Mean[3]
1956-65	250	226	16.25	230.90
1966-75	349	231	21.00	234.11
1976-85	515	228	18.00	232.52
1986-95	835	230	18.00	234.50

However, different results are presented in Table 6.2. The median graduation age increased from 24 years for the graduation cohort 1956-65 to 26 years for the last cohort investigated. The mean age at graduation also shows an immense upward shift to approximately 24 months.

[2] $x_{0.75} - x_{0.25}$

[3] Since the distribution is not symmetrical and additionally afflicted with outliers, the moments such as the mean and the variance are not an appropriate measure. The mean serves here just for comparison with the median.

Table 6.2: Age at graduation in years over 10-year graduation cohorts from 1956 to 1995

Graduation cohorts	n	Median	Inter-quartile range[4]	Mean[5]
1956-65	72	289.5	22.75	291.13
1966-75	220	307.0	35.00	309.40
1976-85	324	305.0	35.00	310.55
1986-95	407	314.0	35.00	316.59

Summarizing both tables it can be stated that although the enrollment age remained constant over time, the graduation age increased. This suggests that the duration of studies increased over the periods investigated. This result is confirmed and described in detail in chapter 7.

6.2 Enrollment Age in Dependency on Influence Factors

Tables 6.1 and 6.2 gave a first impression and some general information about variations in enrollment and graduation age over time. The central aim of this section is to find out whether gender, municipality type and secondary education affect them as well.

Figure 6.1 pictures gender disparities regarding the age at the time of enrollment. It is not surprising that females enrolled earlier, as most males do their obligatory military service or alternative community service – which covers a period between eight and twelve months – immediately after school. Thus, considering enrollments, at the age of 19 almost 75% of females but only 61% of males were already enrolled.

Figure 6.1: Enrollment age in dependency on gender for the enrollment years 1966-95

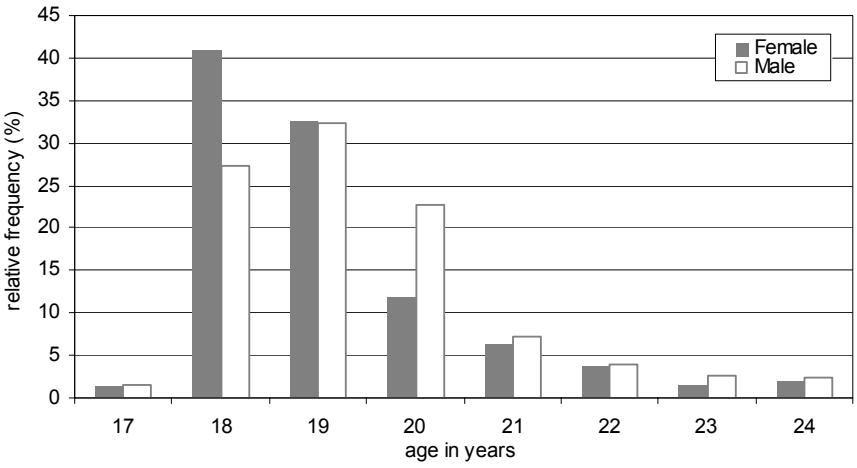


Figure 6.2 displays differences in enrollment age regarding municipality area. At the age of 18 29% of the individuals coming from rural areas, but 38% of those from urban areas are already enrolled. A plausible explanation for this time lag may be the fact that individuals in

[4] $x_{0.75} - x_{0.25}$

[5] Since the distribution is not symmetrical and additionally afflicted with outliers, the moments such as the mean and the variance are not an appropriate measure. The mean serves here just for comparison with the median.

rural areas more often attend vocational and technical colleges, usually lasting one year longer than secondary academic schools, which are more frequented in urban regions.

Figure 6.2: Enrollment age in dependency on municipality type for the enrollment years 1966-95

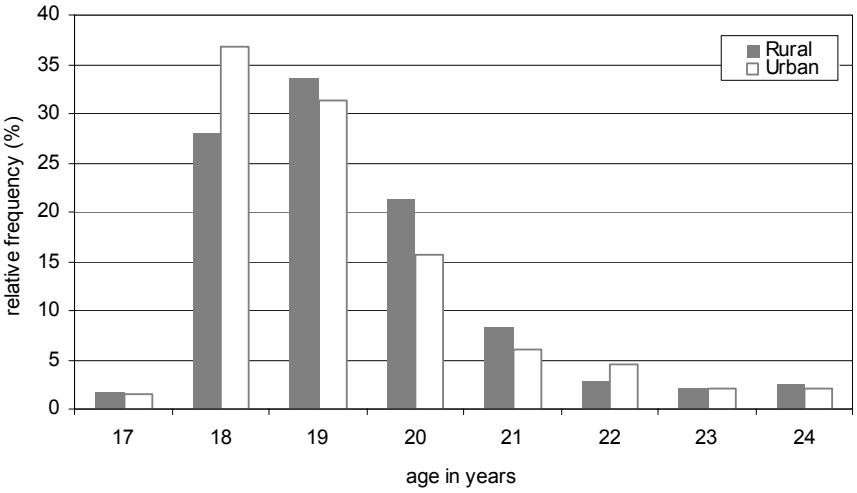
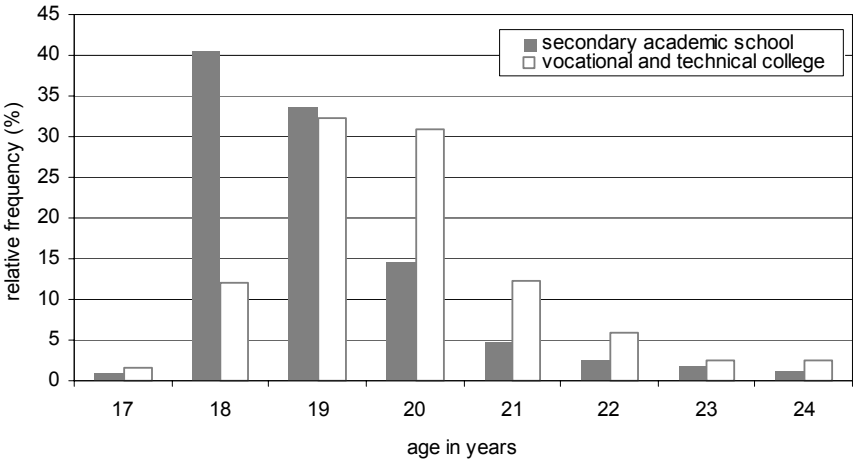


Figure 6.3 shows that the majority of individuals graduating from secondary academic school (40%) immediately enrolled at the age of 18. However, individuals from vocational and technical colleges normally have to attend school until they are 19 and so most of them enrolled at 19 or 20. In contrast to secondary academic school, vocational and technical school imparts full vocational training making it possible to work immediately after graduating. This leads to our hypothesis that vocational and technical college graduates may first enter the labor market and enroll at a later stage in their lives. As a matter of fact, our assumptions are reflected in the data, where one can see that more graduates from vocational and technical colleges than from secondary academic school enrolled later.

Figure 6.3: Enrollment age in dependency on secondary education for the enrollment years 1966-95



6.3 Graduation Age in Dependency on Influence Factors

Figure 6.4 shows that the gender differences found when analyzing enrollment age continue to play a role when considering graduation age. Due to the fact that females start their studies earlier than males, they also finish them earlier. Until the age of 26, in each age group – with the exception of the first one – a higher percentage of females than males graduated. Of all those individuals who graduated between 1966-95, 59% of females but only 44% of males already held a university degree at the age of 25.

Figure 6.4: Graduation age in dependency on gender for the graduation years 1966-95

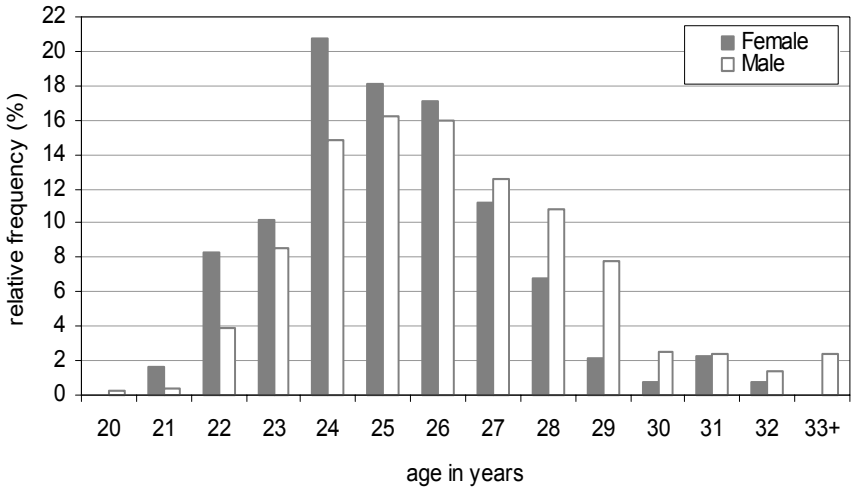


Figure 6.5 shows the graduation age in connection with the type of municipality and indicates that until the age of 24 more individuals from urban areas than from rural regions graduated. The interpretation in connection with Figure 6.2 evidences that the higher graduation age of individuals in rural regions is a consequence of their later enrollment.

Figure 6.5: Gradation age in dependency on municipality type for the graduation years 1966-95

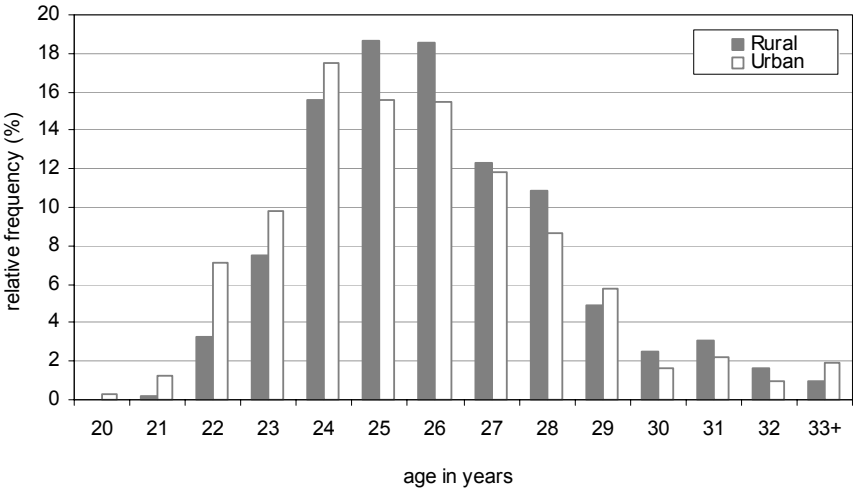
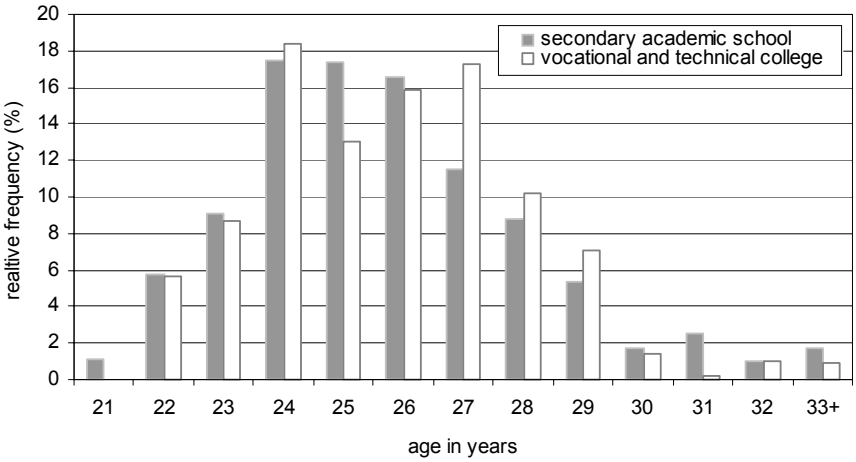


Figure 6.6 demonstrates that more than 70% of individuals who graduated from secondary academic schools or vocational and technical colleges acquired their university degree between 23 and 27.

Figure 6.6: Graduation age in dependency on secondary education for the graduation years 1966-95



7 Duration of University Studies

After having examined the age at the time of enrollment and of graduation, the duration of studies will be looked at more closely as it is of particular interest in the microsimulation. First, relevant changes over time will be identified by analyzing graduation cohorts. Then the factors already applied above will be analyzed to reveal their importance regarding the duration of studies until graduation or dropout.

A central shortcoming of this analysis is the fact that in reality there are other important factors which have a considerably higher influence on the duration of studies. One of them is the fact of whether or not the student is in employment. As it is quite common to work part or even full time during the study, the hours per week employed crucially affect the time used until graduation. As we lack data regarding the employment status of individuals during their university education, we are unable to assess the precise effects. Other aspects that influence the duration of studies and that especially concern females are pregnancy and childbearing. Although we lack data to evaluate the influence of these two aspects, these are interesting research questions that would deserve extra attention.

7.1 Duration until Graduation

In order to discover potential changes over the time span we again consider graduation cohorts. Table 7.1 reveals that the median graduation duration extended by 2 years – from initially 59 to 83 months. While 50% of the individuals of the cohort 1956-65 graduated after five years, this time span lengthened to seven years for the last graduation cohort observed. All these figures confirm the trend towards an extended duration of studies. Again one has to be aware that the sample size decreases enormously, the further one goes back in time.

Table 7.1: Graduation duration in years over 10-years graduation cohorts from 1956 - 1995

graduation cohorts	n	Median	Inter-quartile range	Mean
1956-65	72	59	12	65.00
1966-75	220	71	33	75.20
1976-85	324	71	36	78.41
1986-95	407	83	24	85.06

In this section the duration of studies is investigated for the last three graduation cohorts, namely 1966-75, 1976-85 and 1986-95. Again, the analysis considers the influence of gender, municipality type and secondary education on the duration until graduation.

Figure 7.1 reveals that there are gender differences regarding the duration of studies. After four years of studies around 13% of both female and 12% of male students have graduated.

Figure 7.1: Duration until graduation in years in dependency on gender for the graduation years 1966-95

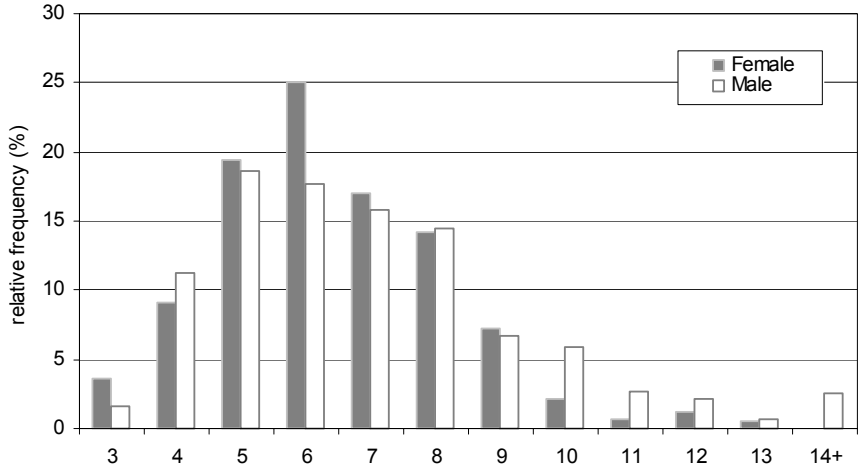
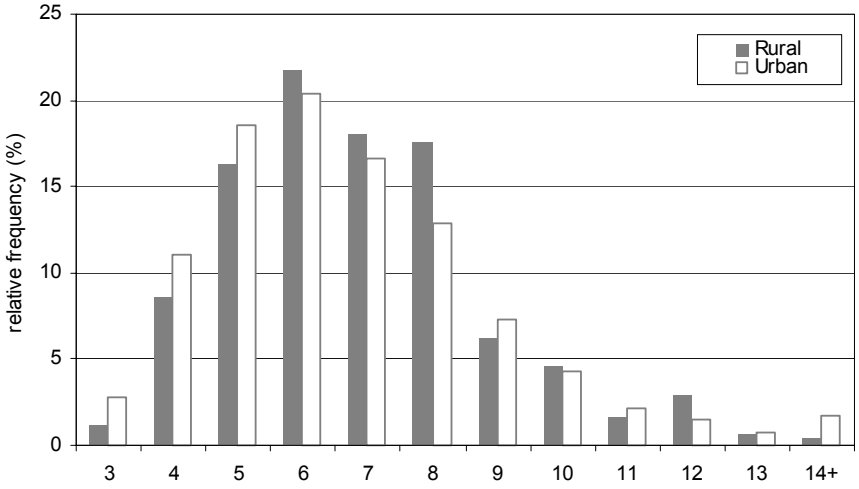


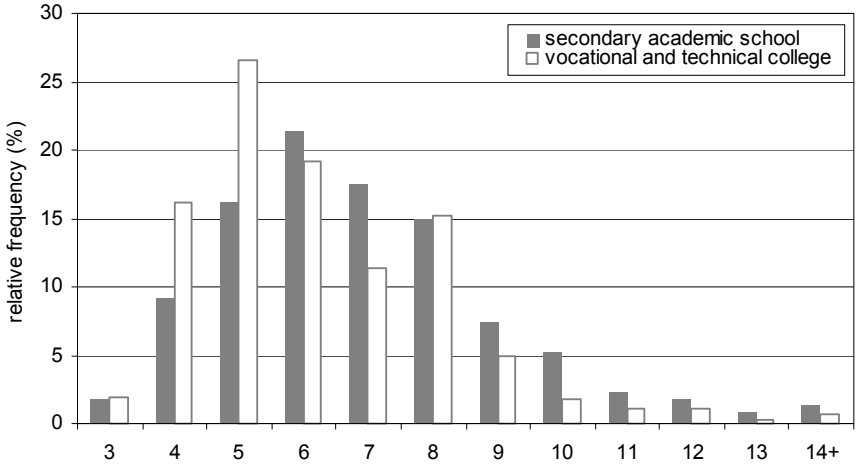
Figure 7.2 shows that students in urban areas obtain their university degree earlier than students living in rural municipalities. When analyzing the picture one has to be aware of the fact that the municipality type refers to the place where people lived at the age of 15. Thus, the fact that most individuals initially living in rural areas move during their studies and live in a dormitory or in a flat at the place of their university is neglected.

Figure 7.2: Duration until graduation in years in dependency on municipality type for the graduation years 1966-95



A completely different and hardly interpretable picture emerges when we want to find out to what extend the previous education influences the length of time until university graduation, as displayed in Figure 7.3.

Figure 7.3: Duration until graduation in years in dependency on secondary education for the graduation years 1965-95



After five years of studies already 44.5% of the graduates of vocational and technical colleges hold a university degree as opposed to 27% of secondary academic school graduates. The main reasons for this may be the fact that depending on their previously attended secondary education, individuals choose distinct type of studies, which vary enormously with respect to their average duration.

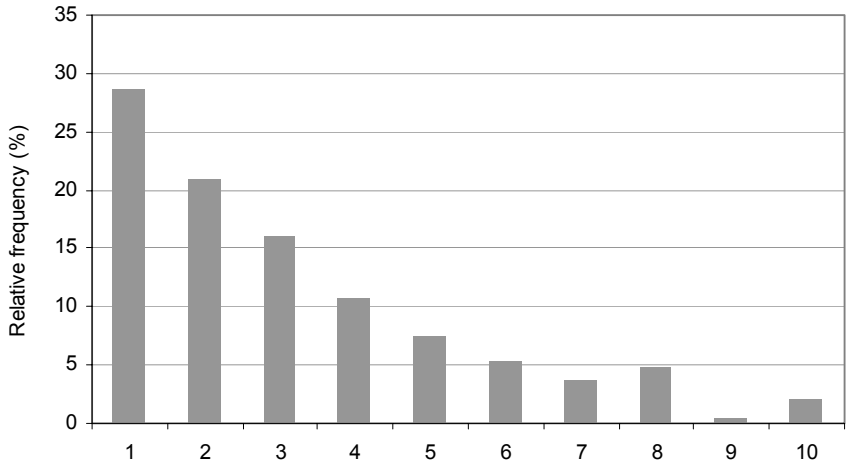
7.2 Duration until Dropout

After having observed the factors by which the duration until graduation is determined, we briefly look at the duration until the point of time when the dropout occurs.

As already mentioned above, we regarded everyone who studied longer than 28 semesters as a dropout in order to avoid having individuals in the sample who enrolled without intending to graduate. Thus, these individuals are comprised in the category 14+.

Figure 7.4 depicts the time interval from enrollment until the dropout date for all individuals leaving university between 1966 and 1995. Of the 255 dropouts in our sample, 29% left university within their first year of studies.

Figure 7.4: Duration until dropout in years for the years of university leaving 1966-95



As opposed to the graduation duration, the time interval from enrollment until the dropout date is much more difficult to attribute to factors like gender, municipality type and previous education, as the dropout is frequently caused by personal reasons. Therefore, we only look at two aspects when investigating the duration until dropout: gender and secondary education. Figure 7.5 makes clear that 62% of female dropouts leave university within the first two years, while this holds true for only 42% of male dropouts. On the other side, notably more females than males are enrolled for more than 14 years and thus are counted as dropouts according to our definition.

Figure 7.5: Duration until dropout in years in dependency on gender for the years of university leaving 1966-95

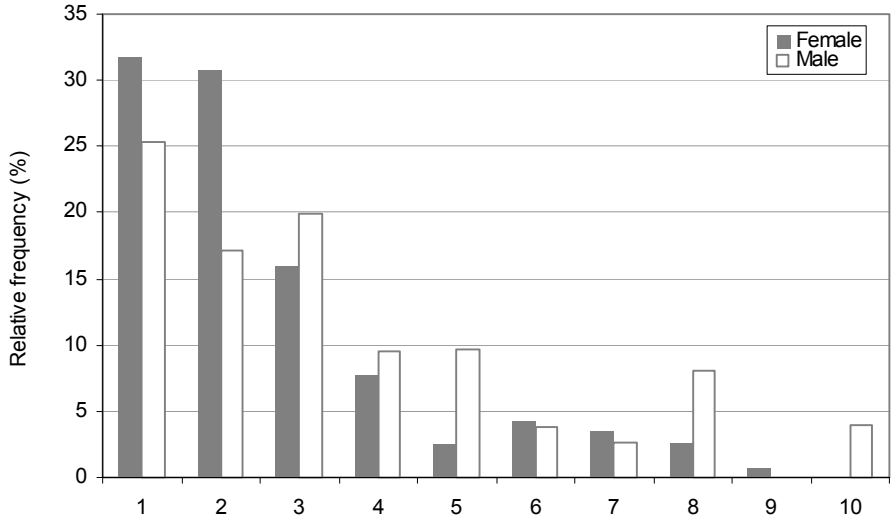
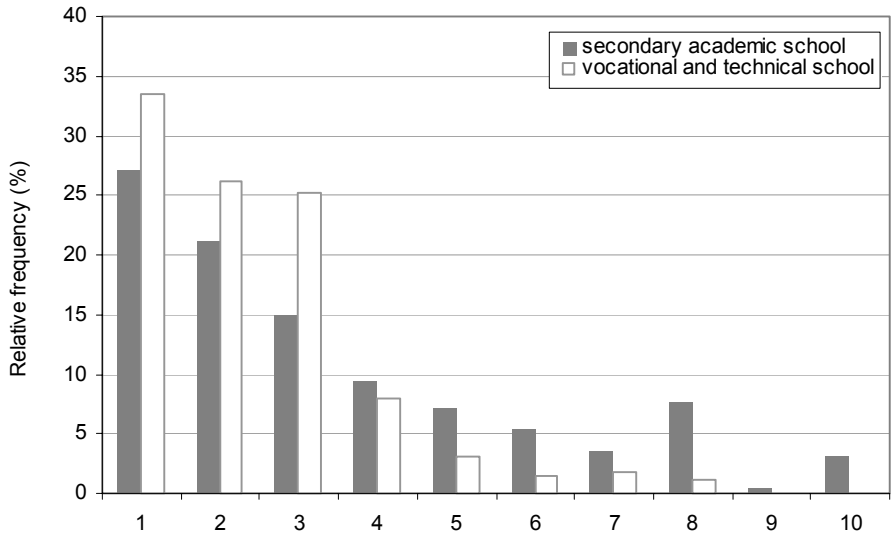


Figure 7.6 shows the dropout duration in dependency on the previous education and yields unambiguous results: Students who acquired their university admission qualification through graduation from vocational and technical school drop out sooner than former pupils of secondary academic schools. 85% of those dropouts who are vocational and technical college graduates left university within the first three years of studies as opposed to 63% of secondary academic school graduates.

Figure 7.6: Duration until dropout in years in dependency on secondary education for the years of university 1966-95



8 Modeling University Education

In our preceding papers we have developed logistic regression models at the transition from primary to lower secondary education (Spielauer et. al., 2002) and from lower to upper secondary education (Schwarz, et. al., 2002). In contrast to primary and secondary education, which is usually completed betweens the ages of 6 and 18, university studies are not completed within a predetermined period or started at a specific age. As the foregoing chapters have shown the duration for completing university studies varies from 3 to 10 years or more, and the enrollment age differs strongly as well. Additionally, some individuals decide to work after finishing upper secondary education and choose to enroll at university at a future stage. In such a case, the admission qualification for university enrollment is often reached via a second-chance education, which can be a *Matura* evening school, a 'Berufsreifeprüfung' or a 'Studienberechtigungsprüfung'. Because of the effect on the pattern of life, such as childbearing, we are interested in modeling first-chance education. Thus, individuals who studied in second-chance education are not of interest in our analysis; nevertheless, the modeling of second-chance education will be considered at a later date. For that reason, we restricted university enrollment age to 24 years. This allows a longer duration in upper secondary education, if for instance individuals attend several types of upper secondary education, as well as the fact that males have to fulfill compulsory military service.

Since not only the background plays an essential role, for instance the education of parents, but the prior life course as well, in the Family Microsimulation Model the previous educational path is taken into account too. In general, there are two ways to enter university education: from secondary vocational and technical colleges (BHS) and from secondary academic schools (AHS), both completed with a *Matura* degree. The reason to separate *Matura* into the two different school types lies in the completely different objective of these schools. While secondary vocational and technical colleges provide individuals with vocational knowledge, secondary academic schools prepare students especially for university education. Hence, for academic school graduates we can expect considerably higher probabilities for university enrollments. Consequently, logistic regression analysis delivers for both admission options a regression equation, which enables us to calculate the transition probabilities from AHS or BHS respectively to university for any combination of the influence factors *education of parents*, *residential area* and *gender*.

Since the university education module of the FAMSIM+ Project is designed of several steps, first, we estimate the probability of university enrollment. Second, the risk of dropping out is estimated. Third, we have to determine the duration of the educational process until graduation or dropout respectively. This is done with a survival analysis, where we perform for both outcomes, graduation and dropout, two separate analyses. By doing so, we obtain for every point in time the probability of dropping out or graduating respectively. Subsequently, we are able to decide with a Monte Carlo simulation for every point in time whether the individual drops out or stays enrolled or whether he graduates or stays enrolled.

8.1 Logistic Regression for University Enrollment

For descriptive reasons and in order to detect changes over time, we will look at the probability of university enrollment of individuals with different social backgrounds and regional destinations first. Afterwards we will consider the fact that actually logistic regression equations at the transition from upper secondary education to university enrollment are required. Table 8.1 shows the goodness-of-fit and the regression coefficient for the logits of the probability of university enrollment over 10-year birth cohorts from 1940 to 1969. The Chi-square statistics of the models are highly significant (p-value=0.000 for all cohorts), which shows that the relationship between the dependent variable and the model is probably real and not due to sampling fluctuations. The pseudo R^2s of around 0.3 are good values for a logistic regression; consequently, these variables are good predictors[6].

The regression coefficients give the information of the partial effect of each variable on the transition probability. The redundant category of the explanatory variables work as reference category; consequently the parameters of the reference category are set at zero. From a practical perspective, the most difficult aspect for logit models is interpreting the logit estimates. However, by taking the exponential of the logistic regression coefficient, we obtain the odds ratio

$$OR = \exp(b_i) = \frac{p_1/(1-p_1)}{p_0/(1-p_0)},$$

which explains the partial effect in the response probability from changing a binary explanatory variable x_i from one to zero, holding all other variables fixed[7]. The odds ratio is usually used for comparison of the proportions in two groups, where the chances of these groups are opposed. If, for instance, for group one the chance of failure is 3:1 and for group two it is 1:5, the odds ratio would be 3:1/1:5, which is 15. Therefore, for group one the chance of failure is 15 times higher than for group two[8]. If the odds ratio is 1, then it is equal for both groups.

The regression coefficients in relation to the odds ratios, listed in Table 8.1, explain the influence of gender, residential area and education of parents on university enrollment for 10-year cohorts from 1940 to 1969. All of the regression coefficients are statistically significant at 95% confidence, except for educating of parents (apprentice) in the cohort of 1940-1949.

[6] For the logistic regression several pseudo R-squares were developed, as equivalents of the usual coefficient of determination. Although the R^2 of the logistic regression is calculated in a different way, from the R^2 in the ordinary linear regression, it can to a great extent be interpreted in the same way.

[7] For derivation and detailed explication of the odds ratio see appendix.

[8] It is important to note that a ratio in the chances of 15 does NOT mean that the probability of one group is 15 times higher than for the other group since OR = $[p_1/(1-p_1)]/[p_0/(1-p_0)]$ and not p_1/p_0, which is a common mistake.

Table 8.1: Logit estimates for the probability of university enrollment over 10-year birth cohorts from 1930 to 1969

Independent variables	Dependent variable: *University enrollment*		
	Birth cohorts		
	1940-49 (n=4737)	1950-59 (n=5425)	1960-69 (n=6025)
GENDER (ref. *female)*			
male	1.218 (0.157) [3.382]	0.691 (0.126) [1.995]	0.402 (0.099 [1.495]
RESIDENTIAL AREA (ref. *rural)*			
urban	0.745 (0.160) [2.106]	0.860 (0.131) [2.264]	0.635 (0.103) [1.887]
EDUCATION PARENTS (ref. *compulsory)*			
apprentice	0.42 (0.224) [1.523]	0.590 (0.183) [1.803]	0.716 (0.150) [2.046]
vocational school	1.646 (0.279) [5.188]	1.672 (0.212) [5.324]	1.791 (0.175) [5.996]
Matura	2.255 (0.223) [9.531]	2.269 (0.189) [9.674]	2.552 (0.159) [12.837]
university	3.416 (0.217) [30.436]	3.486 (0.197) [32.647]	3.311 (0.171) [27.418]
intercept	-4.975	-4.422	-3.910
Chi-square	484.26 (df = 6)	613.70 (df = 6)	808.69 (df = 6)
Pseudo R^2	0.290	0.280	0.272

(Standard error)
[Odds ratio exp(b)]

For males an advantage in the past can be observed; the situation for females improved until the difference between males and females became only minor. In 1940 to 1949 the probability of university enrollment was 3.4 times higher for males than for females, but in 1960 to 1969 the odds ratio was reduced to 1.5. We can also detect a moderate influence of the residential area. However, changes over time are negligible in this connection.

For the five categories of the variable education of parents dummy variables were built, whereby the last category, of course, is linear dependent to the rest of the categories. Therefore, the regression coefficients show not only the influence of education of parents on university enrollment but also the distance to the reference category. In our preceding paper about lower and upper secondary education, the positive influence of the higher educational background of individuals' parents are for individuals of parents with lower education much smaller and over time no improvement can be seen, i.e. the probability of enrolling at university for individuals of parents with university education is around 30 times higher than for individuals of parents with compulsory education.

The fact that the number of students increased considerably the last decades is the result of the educational expansion in society. More and more parents who have a better education also enable their children to receive a good education. However, the individuals of parents with lower education remain strongly discriminated and changes over time are negligible.

Table 8.2: Logistic regression at the transition from previous education to university for individuals born 1950 to 1969

Independent variables	Dependent variable: *University enrollment*	
	Academic school (AHS) to University (n=1257)	Vocational college (BHS) to University (n=958)
GENDER (ref. *female)*		
male	0.721 (0.121) [2.056]	0.754 (0.175) [2.126]
RESIDENTIAL AREA (ref. *rural)*		
urban	0.453 (0.125) [1.574]	0.376 (0.175) [1.456]
EDUCATION PARENTS (ref. *compulsory)*		
apprentice	0.272 (0.182) [1.313]	0.460 (0.240) [1.584]
vocational	0.737 (0.218) [2.089]	1.066 (0.273) [2.904]
Matura	0.861 (0.181) [2.367]	0.905 (0.273) [2.473]
university	1.475 (0.191) [4.373]	1.567 (0.329) [4.793]
intercept	-1.151	-2.607
Chi-square	149.9 (df = 6)	69.74 (df = 6)
Pseudo R^2	0.150	0.111

(Standard error)
[Odds ratio exp(b)]

Table 8.2 shows the goodness-of-fit and the logit estimates for the transitions from upper secondary education to university enrollment both for the transition from upper secondary academic school (AHS) as well as from vocational and technical colleges (BHS). In the analysis we included individuals born between 1950 and 1969, as we were interested in the most current situation but excluded individuals who have not yet finished their education. Again, all of the regression coefficients are statistically significant at 95% confidence, except for educating of parents (apprentice) at both transition states.. The Chi-square statistics of the models are highly significant (p-value=0.000). The pseudo R^2s of 0.150 and 0.111 respectively are moderate values for the logistic regression; consequently, these variables are

moderately good predictors. However, since our interest lies particularly in the estimation of the effect of the x_i on the response probabilities $\Pr(y = 1|x_1, x_2,\ldots x_k)$, the goodness-of-fit is not as important as statistical significance of the explanatory variables.

When comparing Tables 8.1 and 8.2 the influence of previous education becomes obvious. Once an individual has chosen an educational path, such as 'AHS', the education of parents plays a less important role, although still affects the decision of university enrollment. Surprisingly, at the transition from vocational and technical colleges to university enrollment, the odds ratios show slightly better chances for individuals whose parents have vocational school than for individuals whose parents hold a *Matura* degree.

In the Family Microsimulation FAMSIM we require the transition rates from one status to the next for each individual with certain characteristics in relation to any combination of the explanatory variables. We obtain the estimated transition probabilities

$$\hat{p} = \frac{\exp\left(b_0 + \sum_{i=1}^{k} b_i \cdot x_i \right)}{1 + \exp\left(b_0 + \sum_{i=1}^{k} b_i \cdot x_i \right)},$$

where b_i denotes the regression coefficients, b_0 is the intercept, and x_i stands for the i^{th} influence factor.

Table 8.3 shows all probabilities estimated in this way for any combination of the explanatory variables at the transition from upper secondary education to university. It is noticeable that the transition rates from the secondary academic school are on the whole very high, marginally influenced by the educational level of parents, gender and residential area. At the transition from secondary colleges only individuals whose parents have vocational school or university had a good chance to get into university, especially when they are male.

Table 8.3: Transition probabilities for individuals born between 1950-69

Influence factors			Transition probabilities in % from	
Gender	Residence	Education of parents	secondary academic school to university	vocational and technical college to university
Female	Rural	Compulsory	24.0	6.9
		Apprentice	29.3	10.5
		Voc. school	39.8	17.6
		Matura	42.8	15.4
		University	58.0	26.1
	Urban	Compulsory	33.2	9.7
		Apprentice	39.5	14.5
		Voc. school	51.0	23.8
		Matura	54.1	21.0
		University	68.5	34.0
Male	Rural	Compulsory	39.4	13.6
		Apprentice	46.1	19.9
		Voc. school	57.6	31.3
		Matura	60.6	27.9
		University	74.0	42.9
	Urban	Compulsory	50.6	18.6
		Apprentice	57.3	26.6
		Voc. school	68.1	39.9
		Matura	70.8	36.1
		University	81.7	52.3

8.2 Logistic Regression Analysis for University Leaving

For the logistic regression model of university graduation or dropout respectively, we considered individuals who left university between 1960 and 1995. We calculated three different models; first, we measured gender, residential area and education of parents; second, we investigated the influence of gender residential area and admission qualification, and third, we included all of those influence variables. Statistically significant regression coefficients emerge for gender, where obliviously females drop out more frequently, which may be due to childbearing. Additionally, BHS graduates leave university without a degree more often than AHS graduates. However, we have to consider that vocational and technical college graduates already hold a final professional degree, whereas academic school graduates depend on further training to receive professional qualification. All other factors seem not to have any influence on the dropout rate. Table 8.5 shows the dropout rates for model 2 of Table 8.4.

Table 8.4: Logistic regression analyses for dropouts[9]

Independent variables	Dependent variable: *University leaving (dropout* = 1*)*		
	Model 1 (n=1102)	Model 2 (n=1141)	Model 3 (n=1060)
GENDER (ref. *female)*			
male	-0.540** (0.149) [0.583]	-0.542** (0.146) [0.581]	-0.553** (0.154) [0.575]
RESIDENTIAL AREA (ref. *rural)*			
urban	0.144 (0.156) [1.155]	0.153 (0.152) [1.165]	0.184 (0.161) [1.202]
EDUCATION PARENTS (ref. *compulsory)*			
apprentice	0.268 (0.242) [1.308]		-0.205 (0.252) [1.228]
vocational	0.077 (0.279) [1.080]		-0.016 (0.288) [0.984]
Matura	-0.013 (0.237) [1.013]		-0.007 (0.244) [1.007]
university	-0.218 (0.239) [0.804]		-0.184 (0.247) [0.832]
PREVIOUS EDUCATION (ref. AHS)			
BHS		0.620** (0.169) [1.859]	0.541** (0.178) [1.718]
intercept	-1.046	-1.202	-1.1180
Chi-square	17.47** (df = 6)	25.44** (df = 3)	24.86** (df = 7)
Pseudo R^2	0.024	0.034	0.035

(Standard error)
[Odds ratio exp(b)]
* statistically significant $p < 0.05$
** statistically significant $p < 0.01$

[9] Individual who left university 1960 to 1995

Table 8.5: Observed and predicted drop out rates for model 2

Influence factors			Drop out rates in %	
Gender	Residence	Admission qualification	observed	predicted
Female	Rural	AHS	26.4	23.1
		BHS	52.8	35.8
	Urban	AHS	22.8	25.9
		BHS	33.3	39.4
Male	Rural	AHS	11.0	14.9
		BHS	20.6	24.5
	Urban	AHS	20.3	16.9
		BHS	26.1	27.5

8.3 Risk of University Leaving at a Given Point of Time

In section 6 of our paper we considered the density distribution of the duration until dropping out or graduating respectively. The risk of dropping out at a given year of studies t_j, conditional on having survived the years before, is of our interest. For the survival analysis we require the risk of failure at a given point of time t_j, given survival at t_{j-1}

$$q(t_j) = \Pr\{\text{failure in } (t_{j-1}, t_j] \mid \text{survival at } t_{j-1}\} \qquad 0 \le q(t) \le 1$$

We obtain the risk of failing q_j at a given point of time t_j by dividing the number of individuals who failed d_j at time t_j by the number of individuals at risk r_j at time t_j:

$$q(t_j) = \frac{d_j}{r_j}.$$

The number of censored observation c_j at time (t_{j-1}, t_j), i.e. individuals who left observation, must be considered as well. Hence, the remaining number of individuals at risk at time t_i is

$$r_j = r_{j-1} - d_j - c_j.$$

In most studies the probability of surviving beyond time t is even more important surviving a specific length of time

$$S(t) = \Pr(T > t).$$

Since

$$\Pr(\text{did not fail in } t_j) = 1 - \Pr\{\text{failure in } (t_{j-1}, t_j] \mid \text{survival at } t_{j-1}\},$$

we obtain the survival probability

$$S(t) = \Pr(\text{survive beyond } t_{j-1}) \cdot \Pr(\text{did not fail in } t_j),$$

which is mathematically denoted

$$S(t_j) = \Pr(T > t) = S(t_{j-1}) \cdot \left(1 - q(t_j)\right) = S(t_{j-1}) \cdot \left(1 - \frac{d_j}{r_j}\right)$$

This derivation can be done backwards until time t_1, then

$$S(t_j) = (1 - q(t_1)) \cdot (1 - q(t_2)) \cdot \ldots \cdot (1 - q(t_j)) = \prod_{i=1}^{j} \left(\frac{r_i - d_i}{r_i} \right).$$

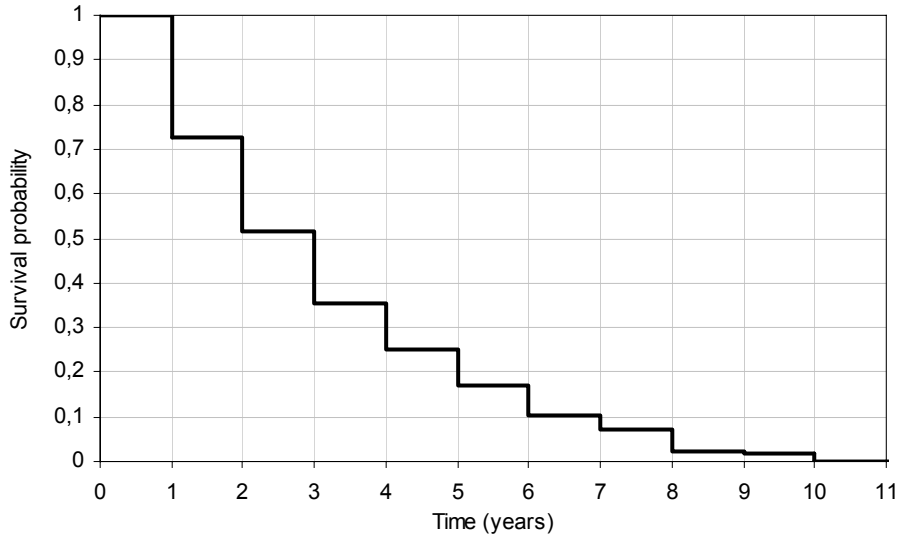
The survival probability $S(t)$ is simply the cumulative product of the conditional survival probabilities. Table 8.6 demonstrates the calculation of the survival probabilities for university dropouts. Since in our model individuals stay at university at least one year respective three years for graduation, and additionally, studies do not take longer than 14 years, individuals can leave university at time t_1 to t_{14}. A result of the following life table that is worth mentioning is that the risk of dropping out $q(t_i)$ of around 0,3 remains nearly constant over all the years.

Table 8.6: Calculation of survival probabilities (Kaplan-Meier survival graph) for university dropouts

Survival time t in years	Individuals at risk r_j	Number of dropouts d_j	Risk of drop out	Survival probability
0	257	0	0	1
1	257	70	0.272	0.728
2	187	54	0.289	0.518
3	133	42	0.316	0.354
4	91	27	0.297	0.249
5	64	20	0.313	0.171
6	44	17	0.386	0.105
7	27	9	0.333	0.070
8	18	12	0.667	0.023
9	6	1	0.167	0.019
10	5	5	1	0

Figure 8.1 illustrates the survival probabilities $S(t)$ of Table 8.6 with a step function. Most of the dropouts occur in the first years of study; afterwards the graph flattens. This is a typical survival function, where the failures are higher in the beginning and reduce over time, such as the function $1 - \exp(-t)$ in the continuous instance.

Figure 8.1: Survival graph for university dropouts



Additionally, we are able to derive the conditional probability directly from the survival probabilities

$$q(t_j) = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})}$$

In our case we possess the relative frequencies of duration until dropout or graduation respectively (cf. Section 6). If we cumulate the density function. we obtain the distribution function

$$F(j) = \sum_{i=1}^{j} p_i \ .$$

Hence, we obtain the survival probabilities directly out of the cumulative frequencies $F(t_i)$

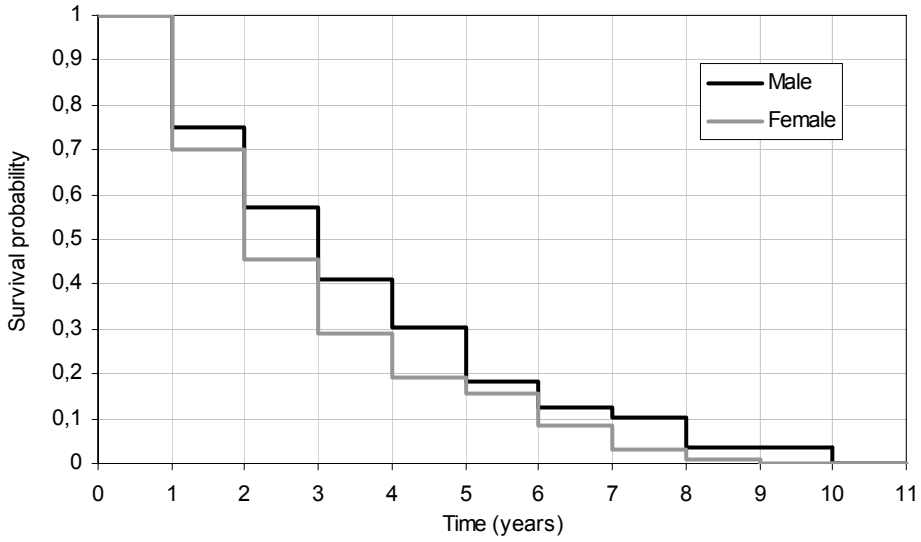
$$S(t_j) = 1 - F(t_j) = 1 - \sum_{i=1}^{j} p_i \ .$$

Consequently, we are able to calculate the risk of failure at a given point of time t_j, conditional on survival the years before $q(t_j)$, straight out of the cumulative frequencies *F(t)*

$$q(t_j) = \frac{F(t_j) - F(t_{j-1})}{1 - F(t_{j-1})} .$$

In view of the fact that we have several influence factors, we might need separate survival graphs for the categories. If significant differences between the graphs occur, the logrank test enables us to compare two or more groups of observations. In doing so, we obtain statistically significant differences for gender only, where females drop out sooner than males, as Figure 8.2 shows.

Figure 8.2: Survival graph for male and female dropouts



Nevertheless, not only dropout but also graduation rates are of interest . For the time until graduation differences occur for the variables gender and previous education. The survival graphs of female and male graduates are shown in Figure 8.3. Again females leave university earlier, yet this time by graduation. Males, on the other hand, seem to study longer. Figure 8.4 shows the differences between individuals from academic schools and vocational

colleges, where presumably BHS provides students with more professional qualifications, so they can graduate earlier from university than students from academic schools.

Figure 8.3: Survival graph for male and female graduates

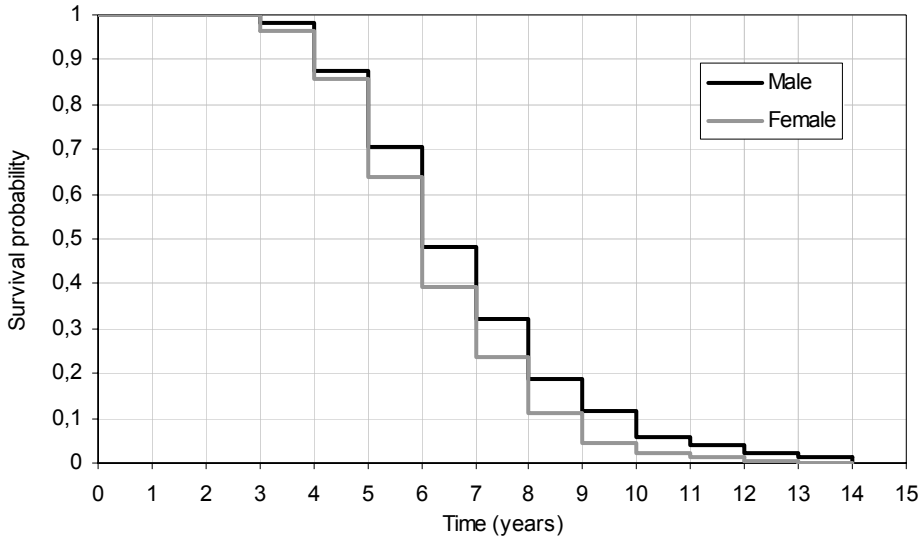
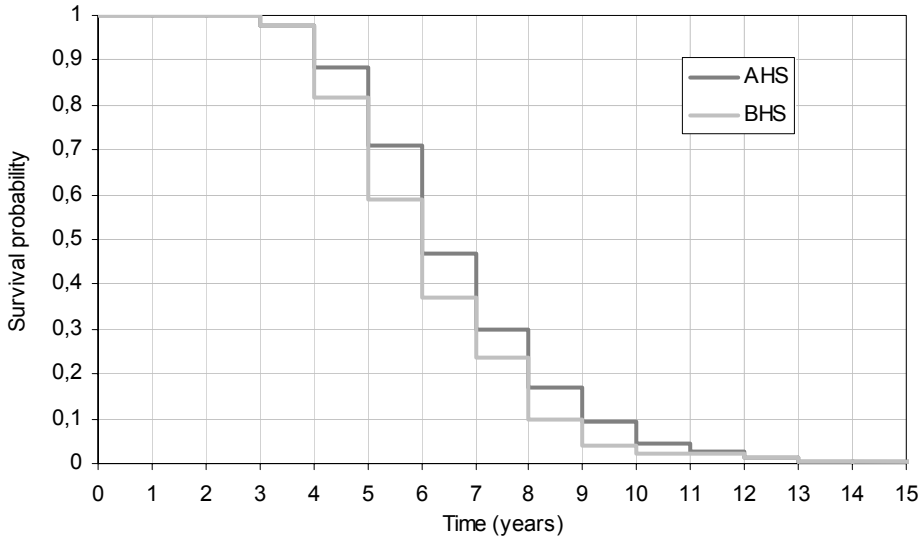


Figure 8.4: Survival graph for university graduates with AHS respective BHS as their previous education



If influence factors in general affect survival, which means that any combination of the factors influence the probability of failing, a survival analysis in dependency on influence factors would be more accurate. This can be done by Cox-regression, also known as proportional hazard regression analysis. However, this would go beyond the scope of this paper. Yet, we will introduce the Cox-regression in one of our next papers.

9 Summary

The findings show that on the macro level a considerable increase of enrollment rates can be observed that more than doubled in the second half of the last century. Female enrollment rates which were only one third of the male rates for the 1935-39 birth cohort have already drawn level with the male rates. While gender differences seem to disappear over time, they can still be found regarding dropout rates that are significantly higher for female students. Another change over time are the long and increasing duration until graduation that, besides the high dropout rates, can be seen as a second typical – and problematic – feature of the Austrian University system. Regarding enrollment duration, some gender differences can be observed: female students both study faster – but also drop out faster. These phenomena on the macro level are the outcome of the decisions of the individual agents on the micro level. Moving from the macro to a micro analysis of university enrollments and graduations, a very stable relation between parents' educational attainments and the university graduation rates of their offspring can be found, besides diminishing gender differences in educational behavior. For given education levels of the parents enrollment rates almost remained constant or even slightly decreased over time for the males, while female rates moderately increased and eventually met the male rates in the last decades. Given the considerable differences in enrollment rates according to parents' education – for the last birth cohort 1965-69 the rates range from around 4% to nearly 60% for the lowest and highest (of the five) education groups – a considerable part of the changes on the macro level can thus be attributed to the changing educational composition of the parents' generation. This result is consistent with the findings regarding prior educational transitions as studied in the previous two papers. Besides that, this strong transmission mechanism already in place at prior educational choices is reinforced at the transition stage to university enrollment. Put differently, the enrollment rates of people already meeting the admission requirements for university vary considerable regarding the parents' education.

Beside the descriptive analysis various regression models and survival graphs regarding university education were developed in this paper. These models will serve as basis of the behavioral modules of the FAMSIM+ model. For university enrollments an advantage for males can be observed in the past; thereafter the situation for females improved, until the difference between males and females became only minor. Between 1940 and 1949 the probability of enrolling university was 3.4 times higher for males than for females, but between 1960 and 1969 the odds ratio was reduced to 1.5. We can also detect a moderate influence of residential area. However, changes over time are negligible. For the five categories of the variable education of parents the positive influence of the higher educational background of individuals parents are much smaller for individuals of parents with lower education and no improvement can be seen over time, i.e. the chances enrolling at university for individuals of parents with university education is around 30 times higher than for individuals of parents with compulsory education only.

Statistically, females drop out more frequently than males, which may be due to childbearing. Additionally, BHS graduates leave university without a degree more often than AHS gradu-

ates. However, we have to consider that the former already hold a final professional degree, while the latter depend on further training to receive professional qualification. All other factors seem not to have any influence on the dropout rate. Most of the dropouts occur in the first years of study; afterwards the graph flattens. Differences in the study duration can be observed between gender, where female graduate but also drop out sooner. Duration differences occur also between the two admission qualification, where individuals from BHS drop out sooner.

The fact that the number of students increased considerably the last decadesis to the result of the educational expansion in society. More and more parents who have a better education also enable their children to receive a good education. However, the individuals of parents with lower education remain strongly discriminated and changes over time are negligible.

Appendix

Logistic Regression for Binary and Multiple Outcomes

The (ordinary) linear regression enables us to predict the value of a continuous variable in relation to one or several explanatory variables. However, in some studies the dependent variable is an indicator for the presence or absence of a condition which can be coded 0 or 1. such as employed/unemployed, success/failure or simply yes/no. In this binary response models we are interested in the probability

$$\Pr(y = 1 \mid x_1, x_2, ..., x_k)$$

For instance. y can be an employment indicator and **x** denotes for instance gender, martial status, education or recent participation in a job-training program. The principle of the logistic regression is similar to the ordinary multiple linear regression. Nevertheless, instead of predicting a value we are able to predict the proportion *p* of individuals with certain characteristics or the probability of a subject having certain characteristics respectively. In order to ensure that the response probabilities are strictly between zero and one, we need a function *G* where

$$\Pr(y = 1 \mid x_1, x_2, ..., x_k) = G(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k) = G(z)$$

where $0 < G(z) < 1$ for all real numbers of *z*.

For a binary outcome variable coded 0 or 1 $\Pr(y = 1)$ is estimated by the proportion of 1s in the sample. If *p* describes the proportion of individuals with certain characteristics then (1 - *p*) is the proportion of individuals who do not have these characteristics. The **odds** $p/(1 - p)$ relate these two proportions and describe the chance. If, for instance, 80% of the patients with a certain disease can be cured, then the chances of success is 0.8/0.2 or 4 to 1. In the logit model one estimates rather a transformation of *p* than *p* itself. If we define *z* as the logit of the probability that an individual will have certain characteristics we obtain the **logit transformation** or **log odds**

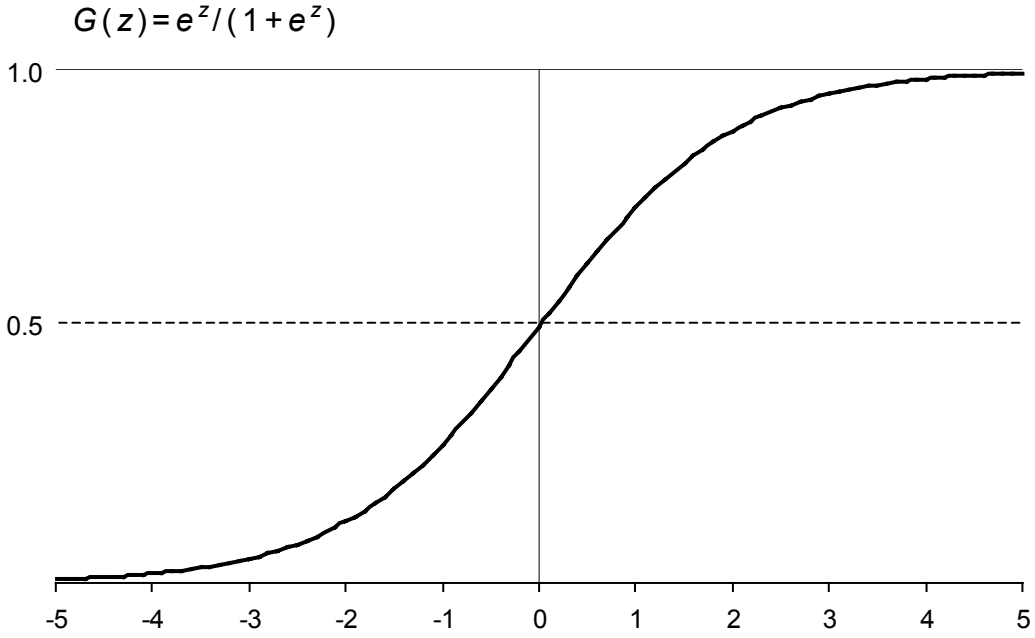
$$z = \text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k .$$

Taking the exponential of the logit transformation and solving for *p* gives us the **logistic function**

$$p = \frac{e^{\text{logit}(p)}}{1 + e^{\text{logit}(p)}} = G(z) .$$

whereby $G(z)$ varies always between zero and one for all real numbers of *z,* as $e^z > 0$. The logistic function is shown in Figure A.1.

Figure A.1: Logistic Function



The approach for multiple outcomes is equivalent: For the l categories of the response variable a binary logistic regression of the $l - 1$ non redundant categories takes place. Thus we obtain the equations

$$\text{logit}(p_1) = \sum_{i=1}^{k} \beta_{ji} \cdot x_i$$

$$\text{logit}(p_2) = \sum_{i=1}^{k} \beta_{2i} \cdot x_i$$

$$\vdots$$

$$\text{logit}(p_{l-1}) = \sum_{i=1}^{k} \beta_{l-1,i} \cdot x_i$$

$$\text{logit}(p_l) = 0$$

For the calculation of the logit estimates β_{ji}. $j = 0 \ldots l$. i = 1…k. the redundant l^{th} category works as reference category; consequently the parameters β_{l0}. β_{l1}. … β_{lk} are set at 0. In the interpretation of the logit estimates. we have to consider this circumstance, as the logit estimates explain the effect of a variable x_i on a category y_j versus the reference category y_l. The transition probabilities of the l^{th} category can be calculated out of the remaining $l - 1$ categories. Thus the transition probabilities for k explanatory variables and l outcomes are

$$p_1 = \frac{e^{\text{logit}(p_1)}}{\sum_{j=1}^{l} e^{\text{logit}(p_j)}}, \quad p_2 = \frac{e^{\text{logit}(p_2)}}{\sum_{j=1}^{l} e^{\text{logit}(p_j)}}, \quad \cdots, \quad p_l = \frac{1}{\sum_{j=1}^{l} e^{\text{logit}(p_j)}}$$

where

$$\text{logit}(p_j) = \sum_{i=1}^{k} \beta_{ji} \cdot x_i \ .$$

In most applications of a binary response model it is essential to explain the effects of the x_i on the response probability $\text{Pr}(y = 1 \mid \mathbf{x})$. The ratio of the odds

$$OR = \frac{\Pr(y = 1 \mid x_i = 1)/\Pr(y = 0 \mid x_i = 1)}{\Pr(y = 1 \mid x_i = 0)/\Pr(y = 0 \mid x_i = 0)} = \frac{p_1/(1-p_1)}{p_0/(1-p_0)}.$$

called the **odds ratio** can be a very helpful tool. The odds ratio is usually used for comparison of the proportions in two groups, where the chances of these groups are opposed. If i.e. for group one the chance of failure is 3:1 and for group two it is 1:5 the odds ratio would be 3:1/1:5, which is 15. Therefore, for group one the chance of failure is 15 times higher than for group two. The odds ratio can be derived directly from the logistic regression coefficient, since the logarithm of the odds ratio gives us

$$\ln(OR) = \ln\left(\frac{p_1/(1-p_1)}{p_0/(1-p_0)}\right) = \ln\left(p_1/(1-p_1)\right) - \ln\left(p_0/(1-p_0)\right) = \text{logit}(p_1) - \text{logit}(p_0)$$
$$= z_1 - z_0 = \beta_0 + \beta_1 x_1 + \ldots + \beta_i \cdot 1 + \ldots + \beta_k x_k - (\beta_0 + \beta_1 x_1 + \ldots + \beta_i \cdot 0 + \ldots + \beta_k x_k) = \beta_i.$$

Accordingly, the exponential of the logistic regression coefficient delivers the odds ratio

$$\exp(\beta_i) = \frac{p_1/(1-p_1)}{p_0/(1-p_0)} = OR.$$

which simply explains the partial effect from changing a binary explanatory variable x_i from one to zero, holding all other variables fixed.

References

Altman, Douglas G. (1990): Practical Statistics for Medical Research. Bocca Raton, London, New York, Washington D.C.: Chapman & Hall/CRC.

Klevmarken, A. (1997): Behavioral modeling in micro simulation models. Department of Economics. Uppsala University.

Lutz, W. E. (1997): FAMSIM-Austria. Feasibility Study for a Dynamic Microsimulation Model for Projections and the Evaluation of Family Policies based on the European Family and Fertility Survey. Schriftenreihe des ÖIF.

Lutz, W., Vaupel J.W. and Ahlburg. D.A. Eds. (1999): Frontiers of Population Forecasting. Population and Development Review. New York: Population Council. 1999. 24 (Supplement).

Neuwirth, Norbert and Spielauer, Martin (2001): Family Microsimulation. ÖIF Working paper 11-2001.

Schwarz, Franz et. al. (2002): Gender and Social Differences at the Transition from Lower to Upper Secondary Education. ÖIF Working paper 23-2002.

Spielauer, Martin (2002): The Potential of Dynamic Microsimulation in Family Studies: A Review and Some Lessons for FAMSIM+. ÖIF Working paper 18-2002.

Spielauer, Martin et. al. (2002): Education and the Importance of the First Educational Choice in the Context of the FAMSIM+ Family Microsimulation Model for Austria. ÖIF Working paper 15-2002.

Recently published WORKING PAPERS:

- Wolfgang Lutz, Vera Nowak, Stefan Ohmacht, Martin Oppitz, Christiane Pfeiffer, Helmuth Schattovits, Martin Schneider, Wilfried Schönbeck, Wolfgang Schulz, **"Rohbericht – Band 1 Ergebnisbericht zum Projekt"**, **"Entwicklung von Modellen eines Kinderbetreuungsschecks und Analyse der Auswirkungen" (Machbarkeitsstudie Kinderbetreuungsscheck)**, Nr. 8/1998 – in deutscher Sprache
- Vera Nowak, Rudolf Schipfer, **"Familien in Zahlen"**, Daten und Graphiken zur Familie in Österreich auf einen Blick, Nr. 9/1998 – in deutscher Sprache
- Paloma Fernández de la Hoz, Johannes Pflegerl, **"MigrantInnen im Spital"**, Workshop für MitarbeiterInnen in Gesundheitseinrichtungen der Stadt Wien, Nr. 10/2001 – in deutscher Sprache
- Martin Spielauer, Norbert Neuwirth, **"Family Microsimulation"**, Nr. 11/2001 – in englischer Sprache
- Veronika Gössweiner, Christiane Pfeiffer, Rudolf Richter, **"Quality of Life and Social Quality"**, Nr. 12/2001 – in englischer Sprache
- Brigitte Cizek, Christiane Pfeiffer, **"HorseTalks",** Nr. 13/2001 – in deutscher Sprache
- Martin Spielauer, Franz Schwarz, Kurt Schmid, **„Education and the Importance of the First Educational Choice in the Context of the FAMSIM+ Family Microsimulation Model for Austria"**, Nr. 15/2002 – in englischer Sprache
- Coomaren P. Vencatasawmy, **"Modelling Fertility in a Life Course Context: Some Issues"**, Nr. 16/2002 – in englischer Sprache
- Norbert Neuwirth, **"Labor Supply of the Family – an Optimizing Behavior Approach to Microsimulation"**, Nr. 17/2002 – in englischer Sprache
- Martin Spielauer, **"The Potential of Dynamic Microsimulation in Family Studies: A Review and Some Lessons for FAMSIM+"**, Nr. 18/2002 – in englischer Sprache
- Sabine Buchebner-Ferstl, **„Die Partnerschaft als Ressource bei kritischen Lebensereignissen am Beispiel der Pensionierung",** Nr. 19/2002 – in deutscher Sprache
- Sonja Dörfler, Karin Städtner, **"European Family Policy Database – Draft Manual"**, Nr. 20/2002 – in englischer Sprache
- Johannes Pflegerl, **"Family and Migration. Research developments in Europe: A general overview"**, Nr. 21/2002 – in englischer Sprache
- Sonja Dörfler, **"Familienpolitische Maßnahmen zum Leistungsausgleich für Kinderbetreuung - ein Europavergleich",** Nr. 22/2002 – in deutscher Sprache
- Franz Schwarz, Martin Spielauer, Karin Städtner, **"Gender, Regional and Social Differences at the** Transition **from Lower to Upper Secondary Education. An Analysis in the Context of the FAMSIM+ Family Microsimulation Model for Austria"**, Nr. 23/2002 – in englischer Sprache
- Veronika Pfeiffer-Gössweiner, Johannes Pflegerl, **"Migration in the European Union: An Overview of EU Documents and Organisations Focusing on Migration"**, Nr. 24/2002/E – in englischer Sprache
- Karin Städtner, **„Arbeitsmarktrelevante Konsequenzen der Inanspruchnahme von Elternkarenz"**, **Nr**. 25/2002 – in deutscher Sprache
- Franz Schwarz, Martin Spielauer, **„The Composition of Couples According to Education and Age. An Analysis in the Context of the FAMSIM+ Family Microsimulation Model for Austria",** Nr. 26/2002 – in englischer Sprache

Order at: Österreichisches Institut für Familienforschung (ÖIF)
Gonzagagasse 19/8, A-1010 Wien
Tel: +43-1-535 14 54-19
Fax: +43-1-535 14 55
Mail: edeltraud.puerk@oif.ac.at

ÖIF

Das Österreichische Institut für Familienforschung (ÖIF) ist ein unabhängiges, gemeinnütziges Institut zur interdisziplinären wissenschaftlichen und anwendungsbezogenen Erforschung und Darstellung der Vielfalt und Veränderungen familialer Lebenswelten aus Sicht von Kindern, Frauen und Männern.

Gedruckt mit Unterstützung des Bundesministeriums für soziale Sicherheit und Generationen sowie der Länder Burgenland, Niederösterreich, Oberösterreich, Salzburg, Tirol, Vorarlberg und Wien.