Lifelong learning inequality? The relevance of family background for on-the-job training

Manfred Antoni*

November 5, 2010

Abstract

Despite ample evidence on intergenerational persistence of formal education as well as on the determinants of non-formal training, these issues have not yet been analysed jointly. The question remains whether people from low-qualified family backgrounds make up for their relatively sparse formal education by investing in non-formal training during adulthood. Hypotheses based on economic and psychological literature suggest otherwise. I use German survey data to estimate the influence of family background on non-formal training participation. Count data analyses show that family background plays a significant role for the likelihood and frequency of on-the-job training. This result holds when controlling for education, ability, personality and job characteristics.

JEL classification: C25, I21, J24, J62

Keywords: on-the-job training, human capital, intergenerational transmission, personality traits, cognitive skills, zero-inflated negative binomial regression

*Institute for Employment Research, Regensburger Strasse 104, 90478 Nuremberg, Germany, email: manfred.antoni@iab.de

1 Motivation

Intergenerational transmission of educational chances and success is documented by an overwhelming number of studies. The educational achievements of children, youths and even adolescents correlate heavily with the success of their parents in this field. Children of better educated parents achieve higher schooling certificates and attempt as well as attain formal occupational training or higher education more often than children from lower social backgrounds. That holds for various countries and has been a stable phenomenon for decades (e.g. Chevalier et al., 2009; Heineck and Riphahn, 2009; Hertz et al., 2007). Several studies show that this is not merely a correlation but rather an causal effect which can be decomposed into an direct effect of parental education and an indirect effect of higher ability inherited from more able parents (Black and Devereux, 2010, p. 27).

This relationship is not restricted to early schooling.[1] Decisions on the attainment of formal training sustain intergenerational inheritance of educational levels even when parents may no longer have direct influence on these decisions. Part of this can be explained by path dependency between schooling and subsequent formal occupational or higher training (e.g. Pallas, 2004). Unequal chances in schooling attainment lead to even more unequal chances when it comes to education that is directly relevant for labour market success. This holds all the more for countries with highly formalised training systems—like Germany—which erect institutional barriers of entry into formal training.[2] The chances of being able to attain more formal education are determined by previous schooling and training attainment. Thus, unequal chances in early education have a lasting influence on educational chances during the whole life-course.

Education—or a lack thereof—heavily influences several aspects of life including labour market success (e.g., Trostel et al., 2002) and a wealth of non-monetary outcomes (e.g., Grossman, 2006). Policymakers therefore strive to minimise inequality in educational chances. Among others, non-formal training, which is henceforth termed as on-the-job training interchangeably, is regarded as a means to compensate for lacking formal educational achievements—hence the term lifelong learning (e.g., European Commission, 2000). Formal and non-formal training[3] are highly dissimilar in terms of, for instance, the amount of time or money they exact, the extent of knowledge and signalling value they offer, or the entry barriers they present. Nevertheless, non-formal training has proven to increase wages, lower the risk of unemployment and facilitate career advancement (e.g., Asplund, 2005; Büchel and Pannenberg, 2004; Dieckhoff, 2007). This

[1] Different types of learning activities will be termed according to their labour market relevance. General education will be termed as schooling, labour market oriented education as training. The term education will include both categories.

[2] See Frick et al. (2007) for a description of the German educational system.

[3] Formal education henceforth denotes activities at educational institutions that provide access to recognised occupational certificates. These are vocational training or higher education. Non-formal activities are courses that take place in and are offered by a variety of institutions but do not lead to recognised certificates (see European Commission, 2000, p. 8).

compensates for lower chances of access to formal education to some extent. People from less educated family backgrounds could catch up in terms of educational and subsequently in labour market chances if they only participated enough in non-formal training. However, for this catching-up to happen they would have to participate in on-the-job training more often or more intensively than people from more favourable family backgrounds. Otherwise the gap widens even more due to the path dependency of formal education.

When we turn our attention to adults' participation in on-the-job training, the empirical knowledge appears to be sound. For Germany, for instance, it is established that participants mainly do not have an immigrant background, are better educated, middle-aged and gainfully employed. These results hold in multivariate analyses based on a variety of data sets (see e.g., Büchel and Pannenberg, 2004; Schömann and Leschke, 2008). Despite the ostensibly ample evidence it is startling that the relation of parental education and an adult's own non-formal training participation has never been analysed explicitly. That holds all the more as a lack of early education due to the influence of parents might indeed predetermine a lack of further training in later life. However, the literature on intergenerational transmission of education concentrates on formal education and thus provides no insights into the question at hand.

There are only few studies that even remotely analyse this relationship. Those who do so measure family background poorly or include it in the analysis without interpreting or motivating it theoretically, thereby shedding no light at the question at hand. Pannenberg (2001) analyses returns to on-the-job training in Germany. In a first step of estimating the selection into training he includes the educational level of the father. This is positively related to the offspring's training participation. Buchmann et al. (1999) include the education and labour market status of the father in their analysis of determinants of non-formal training in Switzerland. They find that the father's labour market status plays no role, regardless of the type of training. The father's education, though, is important for training participation as a lack of a vocational certificate on behalf of the father lessens the probability of his offspring's further training. In both studies, the measurement of family background is so undifferentiated that a thorough analysis of this aspect is not feasible.

The goal of this study is not to show a mono-causal link between parental education and their offspring's non-formal education. Instead, we will learn that on-the-job training during adulthood is determined by a wealth of factors, many of which are related to the parental background. Family background taken by itself plays a significant role for on-the-job training. Factors like cognitive skills, personality traits, cultural capital and the detailed formal educational history will allow us to differentiate between the pure influence of the parental education and other inheritable characteristics.

By using count data models, the analysis allows us to differentiate between what determines non-participation and what is important for the rate of on-the-job training participation over the course of employment spells. We will see that not all relevant factors determine both

measures in the same way. Newly available data from the German ALWA-survey with a strong focus on formal and non-formal education as well as a wealth of social background information on respondents make this kind of analysis possible.

Although the results presented here stem from Germany, they are informative from an international perspective for two reasons. First, Germany shares the phenomenon of high intergenerational persistence with a multitude of countries. Lessons on how to tackle this problem should be relevant to educational and labour market policy in just as many countries. Second, as the German vocational training system with its inherent institutional barriers has been a model for similar systems in several countries, they more than others need evidence to evaluate their educational system's capability to decrease the intergenerational inheritance of unequal educational chances.

This paper develops in the following way. Section 2 shows the theoretical foundation and derives some hypotheses. Section 3 introduces the data set and shows descriptive results. The econometric strategy is described in Section 4. This is followed by a presentation of the results in Section 5. Section 6 adds predictions and results from sensitivity analyses. Section 7 concludes.

2 Theory and hypotheses

Since there is no unified theory on how family background should influence a person's educational decisions in adulthood, one has to relate to a combination of theories on different aspects of the proposed relationship. Indeed, there are some theories that can provide the basis for hypotheses as they border at the topic or concentrate on some aspects of it. The framework for the hypotheses will be provided by human capital theory (e.g., Becker, 1962; Schultz, 1961). Additional assumptions based on theories and results from disciplines such as psychology and genetics are implemented below.

According to Becker (1962) human capital denotes a bundle of skills that determine a person's labour market productivity. In the standard model this productivity can be observed accurately by both the person holding it and potential employers at no cost. Human capital can be augmented by training which is therefore turned into an investment in productive capabilities. This in turn increases future income prospects as a person is assumed to achieved a wage according to her marginal productivity.

Such investments come at a cost, i.e. the direct cost of training as well as indirect costs such as opportunity costs due to reduced productivity during the time of training or the loss of leisure time. Moreover one has to take into account that the value of income earned in the future is not equal to that of income earned today, i.e. to calculate the present value of a future income stream one has to discount each payment by a given interest rate. It follows that investment in human capital should take place as long as the expected present value of wage gains at least equals the present value of direct and opportunity costs. Another important

assumption is that the costs of training are negatively related to the innate abilities of the learner.

Becker and Tomes (1979, 1986) bring the family into the theoretical framework by developing a model that allows for utility maximisation over time on the family level. Rather than treating people as isolated beings, they are assumed to be members of families which consist of several generations. The utility of future generations inside this family is considered in consumption and investment decisions of each actual family member. Parents strive to maximise the utility of the family by creating inheritable wealth but also by investing in the human capital of their offspring. Models in a similar fashion or developments have been proposed by Becker and Chiswick (1966), Solon (2004) and Checchi (2006).

Since, for one thing, training costs depend on the ability of the learner and, for another thing, the amount of investments in the offspring's human capital depends on the expected returns, parents are willing to invest more in the education of more able children. Signals of high ability like cognitive skills or schooling success, which are both observable by parents, would lead to higher investments in the children's human capital. Thus, children who perform well in school are more likely to receive financial assistance by their parents during initial formal schooling and even after its end.

Bowles and Gintis (2002) stress that parents do not only provide the financial means for the education of their offspring. They also hand on important endowments to their children—cognitive ability and personality traits. Surveying decades of research, Plomin et al. (2008) show that both types of endowments are genetically inheritable to some extent. In a study on Germany Anger and Heineck (2010) corroborate that for cognitive ability. After birth the development of the offspring's abilities and personality is affected by a wealth of factors—for instance by the parents' education or parenting skills (see e.g., Cunha and Heckman, 2008; Feinstein et al., 2004). In terms of the model of Becker and Tomes (1986) it is not relevant which part or how much of the inherited endowment is due to genetic inheritance or due to imitating and learning from parents. Both channels are equally included in their model.

Ability and personality on the other hand are related to the propensity of attaining education. The survey by Colquitt et al. (2000) on personality traits that influence training participation reveals that locus of control (see Rotter, 1966), anxiety (see McCrae and John, 1992) and self-efficacy (see Bandura, 1994) are the most predictive traits. Moreover, Fouarge et al. (2010) empirically find several personality traits that strongly influence the willingness to participate in on-the-job training. Cawley et al. (2001) and Heckman et al. (2006) give an overview of studies that show the strong positive relationship between cognitive ability and educational attainment. Cognitive skills and personality traits also positively influence wages (e.g., Green and Riddell, 2003; Heineck, 2011; Heineck and Anger, 2010). By reducing liquidity constraints these endowments also foster human capital investments indirectly.

Moreover, ability and personality are interrelated. Results on the relationship between

cognition and temporal discounting indicate that the higher one's abilities, the lower one's discount rate (Dohmen et al., 2010; Frederick, 2005; Kirby et al., 2005). Applying this to human capital theory, which states that the higher the discount rate the lower the probability of investment in human capital, it follows that the higher a person's ability, the lower her discount rate, the more likely will she participate in education.

To sum up, family background both indirectly and directly influences the incidence of non-formal training. Smarter parents are better educated and thus have higher incomes. They are able to invest more in their offspring's formal education. This in turn puts children on a path of continued education in later life. It follows that they have more access to non-formal training in later life just by having enjoyed better initial conditions than people from a less educated family background. Apart from this indirect effect, ability and favourable personality traits inherited from parents directly increase educational attainment. These endowments lead to higher training motivation, a higher willingness to finance training and more learning success. Similar conclusions can be drawn from the sociological literature on the concept of cultural capital (see e.g., Boudon, 1974; Bourdieu, 1986).

However, the decision whether someone is trained and, if she is, how this is financed is not entirely up to the potential learner herself. Contrary to human capital theory, which deals with labour supply and its quality, theories of segmented labour markets (see e.g., Leontaridi, 1998; Taubman and Wachter, 1986) stress the influence of job characteristics. They argue that training participation mainly depends on the job, or more specifically, the labour market segment someone is employed in. Jobs in some segments mostly provide well-paid and stable employment with good career prospects as well as training opportunities. Other segments comprise of badly-paid jobs which lack stability or career chances. Due to the short expected duration of these employment relationships, time to cash-in returns on investments in human capital would be too short. Anticipating this, employers are not willing to finance such investments. Furthermore, as people in secondary segment jobs earn less, their financial means to invest in their human capital are smaller, making training even more unlikely for them.

Along the lines of these theories I argue that a strong influence of job and firm characteristics is to be expected. However, some influence of family background of a worker should prevail. This is due to the impact of family background on early labour market success, which in turn yields some influence on the actual job situation. This implies that some of the influence of job characteristics is in fact spurious because it only shows up due to state dependence in the first place. Thus, the relation between family background and on-the-job training does not represent a causal relationship between parental characteristics and the outcome variables. The family background—measured as parental qualification and occupational status during an individual's adolescence—should rather be interpreted as proxy variables for unobserved inherited characteristics on the part of the offspring. This leads to the hypotheses and implications summarised below.

H1: The better educated and the more economically successful the family background, the more likely is on-the-job training participation (intergenerational persistence hypothesis).

According to the model of Becker and Tomes (1986) better educated and wealthier parents are more able to support their offspring financially. This support has two implications. For one thing, the theory allows for continued financial support even after the end of formal education or after the offspring has moved out of the parents' household. For another thing, early financial transfers in wealthy families may lead to higher accumulated wealth on the part of the offspring. This in turn facilitates the own financing of on-the-job training in the long run. As financial support of that kind or the offspring's wealth are not observed in the data set at hand, its impact can not be distinguished from that of parental education.

H2: Cognitive skills and favourable personality traits positively influence training participation (endowment hypothesis).

The smarter a person, the higher her probability to invest in any kind of education. Basic cognitive skills and on-the-job training should thus be positively related. The likelihood also depends on personality and cultural capital. The higher the cultural capital and the higher the importance of job success, the more likely will on-the-job training be.

H3: The higher the formal education the higher the probability of on-the-job training (path dependency hypothesis).

The individual wage depends on one's productivity which in turn depends on the training one has received. Therefore, the higher one's schooling and training the higher are the financial means to invest in further education. The investment decisions claimed by human capital theory can be realised as they are not hampered by liquidity constraints.

H4: Participation in on-the-job training is the more likely the higher the job requirements, the higher the weekly working and the larger the firm (job segment hypothesis).

As proposed by theories of segmented labour markets, the job one is employed at strongly determines training participation. In particular, the probability of on-the-job training should rise with weekly working time, job requirements as well as firm size. All of those characteristics are seen as proxy variables for stable employment in a favourable labour market segment and the presence of an internal labour market in the firm. This in turn increases the likelihood that human capital investments will pay off for the financing party.

3 Data and descriptives

I use retrospective survey data of the German study "Work and Learning in a Changing World" (ALWA)[4] which provides monthly longitudinal information on educational and labour market histories. The representative sample consists of 10,177 German citizens aged 18 through 50. Detailed information on educational activities allow a clear distinction between formal education and non-formal training. Aided recall techniques were used during the interviews (Drasch and Matthes, 2011) to reduce recall error.

In the longitudinal part of the survey a wealth of information on educational activities has been gathered. This includes the levels of schooling and training achieved over time as well as their timing in the life course. Non-formal training is measured by its frequency during spells of employment, unemployment and other situations. The survey also covered a wide range of questions on job and firm characteristics for any reported employment spell including marginal employment or parallel jobs.

The longitudinal information is complemented by cross-sectional data on the highest level of education and the employment status of both parents. These variables refer to the status when the respondent was 15 years old. The data also include self-reported measures of personality traits and aspirations. Cognition is measured by self-reported schooling success. These indices may seem as rather crude measures of cognitive ability. Nevertheless, these are valid proxy variables because cognitive ability and educational achievement are highly correlated (e.g., Deary et al., 2007; Plomin et al., 2008).

Independent variables

Standard socio-demographic characteristics include age-classes, sex and the presence of an immigrant background. This is supplemented by a dummy variable for being employed in East Germany and the regional unemployment rate (East/West-Germany). Education is measured by dummy variables for acknowledged German schooling and training levels. Employment experience and its square are included as an additional measure for human capital.

Family background is measured by schooling and training levels of the parent of the same sex as the respondent as well as this parent's employment status during the youth of the survey respondent. Along the lines of the sex-role model (e.g., Sinnott, 1994) the parent of the own sex is likely to have a stronger influence on educational or career decisions of the offspring than the parent with the opposite sex. To analyse whether this restriction has any influence on the results Section 6 presents results based on the highest training level of any of the parents.

To test the endowment hypothesis, several self-reported measures of cognitive skills, personality traits and cultural capital are included in the analysis. These have been computed by principal component analyses (see e.g., Jolliffe, 2002) using several 5-point items for each

[4] See Antoni et al. (2010) for more detailed information on the data set. Technical documentation and information on data access are provided at http://fdz.iab.de/en/FDZ_Individual_Data/ALWA.aspx.

score. Cognition is measured by scores in the three domains prose and document literacy as well as numeracy. Cultural capital is measured by the participation in high-cultural activities. Information on personality traits include scores on an external versus an internal locus of control. A high value on, for instance, an external locus of control signifies the belief that events in life or their outcomes are determined by external factors such as luck rather than by own decisions, actions or effort. This belief should lead to a low expected return to education and thus to fewer investment in on-the-job training. Personality is also measured by a score of employment-related self-confidence. Finally, scores on the importance of life domains such as work and occupation on the one hand and family and friends on the other hand are included. I argue that these scores represent proxy variables for time preference as they relate to the trade-off between work and leisure time. The more value one assigns to the family or friends, the higher the time preference and the lower the likelihood to invest in work-related training.

Two potential shortcomings of these scores have to borne in mind. First, the items are self-reported in relatively short item batteries instead of being measured by established psychometric tests. Thus, measurement error due to recall error cannot be ruled out. Second, a crucial assumption underlying these scores is that what has been reported during the interview has been stable for a given respondent in the past and has been relevant during every employment spell in the data. Borghans et al. (2008) review the literature that calls this assumption into question. However, neither this assumption nor the accuracy of the self-reported measures can be tested with the data at hand. This would make repeated and more detailed interviews necessary.

As only spells of dependent employment will be considered, firm and job characteristics control for the employer's influence on human capital investments. The data thus also allows for tests of the implications of theories of segmented labour markets. The data provide information on the working time relative to full-time employment, the job requirement, the firm size and whether the worker is employed in the civil service sector. Sample statistics for all independent variables are given in Table A.1 in the appendix.

The analysis sample considered in Table A.1 has been restricted by exclusion of the following groups: Spells that originate from outside Germany are excluded; spells that originate from East Germany starting before reunification are excluded as their training decisions do not necessarily follow the cost-benefit considerations laid down by human capital theory. Although the same could apply to public sector workers, they are not excluded from the sample. The implications of this decision are presented in Section 6. Finally, self-employment spells are excluded as the notion of co-financing of training by the employer cannot be examined for them. The number of observations lost due to each of these exclusions and to missing values is given in Table A.2 in the appendix.

Although descriptive in nature, Table 1 provides a first empirical indication of whether individual characteristics are associated with family background variables as stated in Section 2.

The table reports on t-tests that compare the means of central individual characteristics by different dichotomous family backgrounds. In doing so, a distinction between the offspring of low qualified and highly qualified family backgrounds is possible. The offspring of parents that lack formal schooling or training is described in the column denoted by "Low", whereas the offspring the parents that have both formal schooling and formal training are described by the column denoted by "High". The table is based on spells rather than individuals because some independent variables can vary between different spells of a given person in the sample.

Table 1: On-the-job training, education and endowments by dichotomous family background (measured as educational level of parent of the same sex as respondent), t-test of difference

	Low (1)	High (2)	Difference	t-statistic
Participation in on-the-job training (dummy)	0.293	0.338	0.045***	(4.985)
Training frequency per spell	1.966	2.405	0.439***	(3.223)
No schooling (dummy)	0.012	0.006	−0.006***	(−3.718)
Lower secondary schooling (dummy)	0.302	0.184	−0.118***	(−15.165)
Intermediate schooling (dummy)	0.439	0.392	−0.046***	(−4.885)
Upper secondary schooling (dummy)	0.247	0.418	0.170***	(18.268)
No training (dummy)	0.058	0.046	−0.012***	(−2.962)
Apprenticeship (dummy)	0.782	0.630	−0.152***	(−16.709)
Master craftsman / technician (dummy)	0.027	0.056	0.030***	(7.010)
Higher education (dummy)	0.134	0.269	0.135***	(16.351)
Prose literacy (score)	−0.002	0.051	0.053***	(2.721)
Document literacy (score)	−0.191	0.081	0.272***	(14.236)
Numeracy (score)	−0.130	0.036	0.165***	(8.741)
High-cultural activity (score)	0.014	0.090	0.076***	(4.099)
Importance of work (score)	−0.135	−0.029	0.106***	(5.498)
Importance of occupation (score)	−0.010	0.149	0.159***	(9.568)
Importance of friends (score)	−0.159	−0.064	0.096***	(4.890)
Importance of family (score)	0.147	0.033	−0.114***	(−6.086)
External locus of control (score)	−0.006	−0.041	−0.035*	(−1.862)
Internal locus of control (score)	0.104	−0.026	−0.130***	(−6.670)
Employment-related self-confidence (score)	−0.000	0.052	0.052***	(2.745)

Notes: ALWA, own unweighted calculations; 17254 observations; ***,**,* denote a significant difference at 1%, 5%, 10%; *Low* denotes a lack of parental formal schooling or training degree; *High* denotes parental formal schooling as well as training.

The first striking difference between columns 1 and 2 is related to the probability of participation in and frequency of on-the-job training per spell. The offspring from more educated family backgrounds is significantly more likely to experience on-the-job training. On average, they also experience on-the-job training more often per employment spell than the offspring from less educated backgrounds, with a frequency of 2.41 compared to 1.97.

The following rows shed some light on how people from different family backgrounds differ in terms of education, cognitive skills and personality. Given the well-documented low educational

mobility in Germany it is hardly surprising that the offspring of better educated parents is better educated itself. This is demonstrated by, for instance, the significantly lower share of observations without formal schooling or training in the second column. Higher cognitive skills, a higher prevalence of high-cultural activities and higher employment-related self-confidence in column 2 also corroborate the assumption of inheritability of cognition and personality. All in all, the offspring of better educated parents seem to be better equipped with endowments and skills relevant for the labour market. Based on the assumptions stated in Section 2 these differences should make people with a better educated family background more likely to experience on-the-job training. The next step will be to analyse this relationship descriptively.

Dependent variable

The number of on-the-job training activities during each employment spell is considered as the dependent variable. Table 2 informs about this variable and differentiates between several individual, parental and job characteristics. The first column reports on the likelihood to experience on-the-job training in a given spell; the second column shows the number of on-the-job training activities per spell. As employment relationships strongly differ in duration column 3 also presents the frequency of on-the-job training per year.

Table 2: Probability and frequency of on-the-job training by individual, parental and job characterstics (only dummy variables)

	Participation (1)	Frequency (2)	Frequency/year (3)	Obs.
Total	0.329	2.320	0.563	17,254
Age: 18-21	0.230	1.639	0.359	4,472
Age: 22-25	0.308	2.190	0.527	3,725
Age: 26-35	0.397	2.974	0.668	6,241
Age: 36-51	0.365	2.125	0.703	2,816
Male	0.330	2.390	0.561	8,713
Immigrant Background	0.319	2.185	0.582	2,812
No schooling	0.145	0.661	0.190	124
Lower secondary schooling	0.226	1.517	0.355	3,562
Intermediate schooling	0.308	2.092	0.444	6,925
Upper secondary schooling	0.410	3.021	0.806	6,643
No training	0.138	0.668	0.279	826
Apprenticeship	0.282	1.896	0.439	11,367
Master craftsman/technician	0.405	2.830	0.493	873
Higher education	0.480	3.693	0.970	4,188
P: no schooling degree	0.223	1.657	0.379	443
P: lower secondary schooling	0.322	2.232	0.484	11,002

(table continued on following page)

	Participation (1)	Frequency (2)	Frequency/year (3)	Obs.
P: intermediate schooling	0.338	2.427	0.678	3,490
P: upper secondary schooling	0.371	2.708	0.800	2,319
P: no vocational degree	0.291	1.910	0.389	3,222
P: apprenticeship	0.329	2.329	0.566	10,641
P: master craftsman/technician	0.368	2.605	0.607	1,447
P: higher education	0.367	2.741	0.804	1,944
P: not employed	0.333	2.134	0.606	2,615
P: employed	0.325	2.327	0.545	12,880
P: self-employed	0.357	2.547	0.630	1,759
Firm in East Germany	0.274	1.999	0.499	2,195
Civil service	0.464	4.118	0.835	3,730
Working .25 full-time	0.145	0.633	0.214	1,429
Working .5 full-time	0.280	1.898	0.537	1,662
Working .75 full-time	0.353	2.672	0.696	931
Working full time	0.354	2.531	0.595	13,232
No training required	0.090	0.349	0.113	1,913
Induction period required	0.169	0.891	0.244	1,983
Vocational training required	0.298	1.917	0.468	8,438
Vocational schooling required	0.468	3.483	0.805	1,100
Master craftsman/technician required	0.453	3.702	0.653	678
Higher education required	0.584	4.802	1.189	3,142
Firm size: less than 5	0.197	1.049	0.322	1,394
Firm size: 5-9	0.243	1.439	0.417	2,222
Firm size: 10-19	0.265	1.613	0.473	2,186
Firm size: 20-99	0.294	1.927	0.478	3,999
Firm size: 100-199	0.330	2.283	0.714	1,830
Firm size: 200-1,999	0.432	3.184	0.711	3,629
Firm size: 2,000 and more	0.472	4.219	0.755	1,994

Notes: ALWA, own unweighted calculations; 17254 observations.

Over the whole estimation sample the probability to experience on-the-job training during an employment relationship equals 32.9%. The mean number of on-the-job training activities per spell is 2.32, whereas the mean frequency per year is 0.56. By considering how these numbers vary by subgroup these descriptive results can provide a first impression of the influence of individual, parental and job characteristics on on-the-job training.

For one thing, the better educated a person the higher is the likelihood of on-the-job training as well as its frequency. This applies for both formal schooling and training levels. For another thing, on-the-job training probability and frequency are associated with parental schooling and training. The better educated the family background the more likely and frequent is own on-the-job training. While, for instance, a parental background without formal schooling

is associated with a probability of on-the-job training of 22.3%, a parental upper secondary schooling degree is associated with a training probability of 37.1%.

The amount of on-the-job training is also related to job and firm characteristics. The higher the weekly working time, the job requirement or the number of employees in the firm, the more likely and frequent is on-the-job training. The descriptive results thus do not contradict the hypotheses given in Section 2. The next section explains the econometric strategy to test these hypotheses more thoroughly.

4 Econometric strategy

The dependent variable under consideration is the number of on-the-job training activities during any given employment spell, denoted as y. Since the underlying random variable is a non-negative integer with a strongly right-skewed distribution, a count data model (e.g., Cameron and Trivedi, 1998; Winkelmann, 2008) is the usual approach for this analysis. Count data models enable us to study the number of events during a measurable amount of time and infer from it the influence of explanatory variables on the frequency of the event using maximum likelihood estimation.

The most basic variant of count data models assumes a probability distribution for the number of events based on the Poisson distribution. The Poisson regression model though can only be justified if the properties of the underlying Poisson distribution are fulfilled, most notably the assumption of equality of conditional mean and conditional variance:

$$Var(y_i|\mathbf{x}_i) = E(y_i|\mathbf{x}_i) = \mu_i \tag{1}$$

If this assumption of equidispersion is violated the single-parameter Poisson distribution may be too inflexible to account for the real-life data at hand. As a result, biased standard errors make statements on the significance of regressors unfeasible. A glance at the distribution of the count variable in the estimation sample indeed reveals significant overdispersion in the data. Its variance of 49.56 is more than 20 times as high as its mean of 2.32.

A common remedy is a generalisation by using a model based on the negative binomial distribution (e.g., Cameron and Trivedi, 1986; Hilbe, 2008). This allows for more flexibility due to an additional distributional parameter (Cameron and Trivedi, 1998, pp. 70 sqq.). The variance function from Equation 1 therefore turns into

$$Var(y_i|\mathbf{x}_i) = \mu_i + \alpha\mu_i^k. \tag{2}$$

One possible interpretation of this gamma-distributed parameter α is that of dealing with unobserved heterogeneity as a mixture distribution. The dispersion parameter α can be estimated from the data. The two most common variants of the negative binomial model are

often called Negbin I ($k = 1$) and Negbin II ($k = 2$). Both are better suited to deal with overdispersed data than the Poisson model but the decision for either of them depends on the data at hand. Information criteria are suitable to provide information for this decision.

Since Poisson and negative binomial models are nested, different tests can be used to determine whether the null hypothesis of α being equal to zero has to be rejected. If that were the case the Poisson regression model would be preferable. Highly significant test statistics of both the Pearson's χ^2 goodness-of-fit test (199229.8, df=17,206) and the likelihood ratio test (86089.4, df=17,206) reject the null hypothesis and reveal that the Poisson regression model is not appropriate for the data at hand. Information criteria show that the Negbin II model (BIC=49338.49, AIC=48958.45) is superior to the Negbin I variant (BIC=49840.55, AIC=49460.52) for the analysis at hand. Due to this result the Poisson and Negbin I regression models will no longer be considered throughout the text.

Both Poisson and negative binomial models produce biased results when the share of zero-counts in the data is much higher than predicted by the underlying probability distribution (Hardin and Hilbe, 2007). The most common solutions in econometrics are the hurdle model (Mullahy, 1986) and the zero-inflated model (Heilbron, 1994; Lambert, 1992). Both models are flexible enough to allow for a substantive amount of non-participants in the sample and can be formulated based either on the Poisson distribution or on the negative binomial distribution. Thus, the issues of missing equidispersion and an excess of zeros can both be tackled by these models.

The models' interpretations of what drives the amount of zero-counts differ. The hurdle model assumes that counts below a given hurdle—in our case zero-counts—and counts above the threshold stem from different data generating processes. They are potentially driven by different explanatory factors. The hurdle model thus is a two-part model with one part modeling the probability to encounter an event at all and a second part considering positive counts of events. The interpretation would be that of a two-staged decision process. Empirical applications of this model to on-the-job training include Arulampalam and Booth (2001) and Pannenberg (1998).

The zero-inflated model enables us to distinguish between a subgroup that is not subject to the risk of any event and a subgroup that can experience any number of events, including zero-counts. Covariates that describe the first group explain the inflation of the number of zero-counts. The covariate vector describing the group that experiences any number of events including zeros can differ from the former. Empirical evidence provides support for this method in the current context. Backes-Gellner et al. (2007) show that some people persistently refrain from on-the-job training as contrasted to those who merely participate less frequently than others. They identify characteristics that make it more likely to belong to the group of chronic non-participants. This is more in line with the interpretation of the zero-inflated model than that of the hurdle model.

Given that the two equations of the zero-inflated model are estimated simultaneously it is generally more efficient than the hurdle model. Bearing that in mind and using both Akaike and Bayesian information criteria to compare the two non-nested models, the hurdle model is rejected in favour of the zero-inflated model. The test based on the work of Vuong (1989) confirms that the data indeed show an excess of zeros as the zero-inflated negative binomial model is preferred over the negative binomial model. Therefore, all results presented in Sections 5 and 6 are based on the zero-inflated negative binomial model and maximum likelihood estimation.[5] For an application of this method see Gerner and Stegmaier (2009) who analyse the provision of on-the-job training by firms in Germany.

The longer an employment spell the more occasions can arise in which on-the-job training is necessary or profitable. Thus, a given count may represent different rates of training per period of time. The natural logarithm of the spell duration in months is included in the model to indicate the amount of time during which a person is exposed to the possibility of experiencing on-the-job training. Its coefficient is constraint to be 1.

Some of the individuals under consideration experience more than one employment spell during the observation period. These observations are correlated due to common unobserved characteristics (Moulton, 1990). This usually leads to an underestimation of standard errors. To achieve robust standard errors clustered on the person-level the variance-covariance matrix is estimated by the modified, cluster-robust version of the sandwich estimator based on Huber (1967) and White (1980).

5 Results

The estimation results given in Tables 3 and 4 are structured according to the succession of the hypotheses given in Section 2. In every model specification, the column concerning the probability never to participate in on-the-job training during the spell is denoted by "Non-participation" whereas the column related to the number of training events per spell is denoted by "Frequency". The former indicates that exponentiated coefficients that are significantly larger than 1 lead to an absence of on-the-job training and therefore inflate the number of zero-counts in the data. When it comes to the column denoted by "Frequency" a value above 1 signifies a covariate that is associated with a higher number of on-the-job training activities per spell. This has to be born in mind when interpreting the exponentiated coefficients. The covariates common to all specifications include age-class, sex, the presence of an immigrant background, employment in East Germany and the unemployment rate.

[5] Details on the tests and the results of all discarded models are available upon request.

Table 3: Determinants of non-participation in training and of the number of courses respectively, zero-inflated negative binomial regression, Probit inflation

	Model 1				Model 2			
	Non-participation		Frequency		Non-participation		Frequency	
Age: 22-25	0.684***	(−3.72)	1.184**	(2.05)	0.777***	(−3.01)	1.184**	(2.04)
Age: 26-35	0.418***	(−3.10)	1.224**	(2.10)	0.598***	(−5.09)	1.204**	(2.10)
Age: 36-51	0.430***	(−3.62)	1.398***	(3.27)	0.629***	(−4.10)	1.378***	(3.06)
Male	1.364	(1.53)	1.017	(0.12)	1.081	(0.85)	1.075	(0.61)
Immigrant Background	1.099	(0.88)	1.093	(0.99)	1.030	(0.33)	1.136	(1.50)
Firm in East Germany	1.520**	(2.01)	0.839	(−1.07)	1.465**	(2.46)	0.792	(−1.47)
Regional unemployment rate	1.014	(0.86)	1.007	(0.43)	1.003	(0.23)	1.018	(1.26)
P: lower secondary schooling	0.649**	(−2.20)	0.919	(−0.46)	0.629***	(−2.83)	0.964	(−0.21)
P: intermediate schooling	0.591**	(−2.30)	1.177	(0.85)	0.577***	(−2.77)	1.158	(0.76)
P: upper secondary schooling	0.498**	(−2.49)	1.182	(0.79)	0.567**	(−2.29)	1.103	(0.46)
P: apprenticeship	0.820*	(−1.68)	1.324**	(2.28)	0.883	(−1.35)	1.284**	(2.33)
P: master craftsman/technician	0.712**	(−2.07)	1.366**	(2.33)	0.818	(−1.42)	1.328**	(2.24)
P: higher education	0.880	(−0.56)	1.570**	(2.48)	1.039	(0.18)	1.509**	(2.34)
P: employed	0.811	(−0.92)	0.812	(−0.73)	0.882	(−0.99)	0.825	(−0.87)
P: self-employed	0.782	(−1.02)	0.893	(−0.39)	0.866	(−0.92)	0.920	(−0.37)
Prose score					0.898***	(−2.84)	1.121***	(3.36)
Document literacy score					0.929**	(−2.38)	1.062*	(1.65)
Numeracy score					0.985	(−0.41)	0.998	(−0.05)
High-cultural activity					0.743***	(−5.28)	1.114***	(2.98)
Importance of work					0.943	(−1.22)	0.943	(−0.75)
Importance of occupation					0.895***	(−3.11)	1.026	(0.55)
Importance of friends					0.960	(−1.04)	1.037	(1.10)
Importance of family					0.941	(−1.57)	0.990	(−0.26)
External locus of control					1.104**	(2.28)	0.958	(−1.29)
Internal locus of control					1.043	(0.87)	1.122**	(2.14)
Employment-related self-confidence					1.054	(1.23)	1.042	(0.85)
Constant	1.237	(0.85)	0.041***	(−11.15)	1.314	(1.25)	0.035***	(−13.04)
ln(α)	2.961***	(6.60)			2.561***	(9.80)		
AIC	50081.5				49619.3			
BIC	50337.4				50045.8			
Wald-Statistic (χ^2)	68.774				170.215			
p-value	0.000				0.000			

Notes: ALWA, own calculations; 17254 observations; 6490 individuals; exponentiated coefficients; z-statistics in parentheses; cluster-robust standard errors; ***, **, * denote significance at 1%, 5%, 10%; reference category: parent of own sex without formal schooling or training and not employed, no own formal schooling or training.

Model 1 presents the most parsimonious specification. It only includes parental characteristics and those socio-demographic covariates that serve as control variables in all specifications. Model 1 corroborates the descriptive results on the relationship between family background and training participation. A better educated family background seems to be related to a decreased likelihood of totally refraining from on-the-job training as indicated by exponentiated coefficients below 1. This relationship is driven by parental schooling as well as their training level. For the number of training events per spell this finding no longer holds. Only the dummy variables representing parental training levels show significant positive influence on the own training frequency. The parental labour market status bears no relevance in either of the two equations. Nevertheless, a Wald-test on joint significance of the background variables shows a significant influence of these variables ($\chi^2(16)$=106.13, p-value=0.00). So far, the intergenerational persistence hypothesis cannot be rejected.

Model 2 extends the specification by including indices based on self-reported measures of

cognitive skills, cultural capital and personality traits. In line with the endowment hypothesis, cognitive skills as well as the extent of participation in high-cultural activities are positively related to both the occurrence and the number of training activities. Self-reported personality traits are only weakly related to on-the-job training. The importance of the own occupation decreases the likelihood of never participating in on-the-job training whereas this likelihood increases with external locus of control. As expected, the stronger the internal locus of control the higher the frequency of on-the-job training. The respective Wald-tests show joint significance of the endowment as well as the family background variables ($\chi^2(22)$=190.79, p-value=0.00 and $\chi^2(16)$=68.22, p-value=0.00, respectively). Both the intergenerational persistence hypothesis and the endowment hypothesis can be maintained.

Additional covariates on educational level and full-time equivalent employment experience prior to the actual spell are included in Model 3 in Table 4. The results support the path dependency hypothesis. Having achieved at least intermediate schooling and any level of vocational degree makes non-formal training abstinence less likely. There is no influence of training levels on the frequency of on-the-job training. A Wald-test can not reject the joint significance of own formal educational levels on on-the-job training ($\chi^2(12)$=94.78, p-value=0.00). The employment experience on the other hand influences the number of training events per spell and hints at a u-shaped relationship. This could be explained by growing opportunity costs due to increasing wages over the life-course.

Table 4: Determinants of non-participation in training and of the number of courses respectively, zero-inflated negative binomial regression, Probit inflation

	Model 3				Model 4			
	Non-participation		Frequency		Non-participation		Frequency	
Age: 22-25	0.888	(−1.29)	1.180**	(1.99)	0.865*	(−1.87)	1.125	(1.45)
Age: 26-35	0.925	(−0.40)	1.251**	(1.98)	0.738***	(−2.74)	1.095	(0.90)
Age: 36-51	1.044	(0.15)	1.477**	(2.40)	0.678**	(−2.48)	1.320*	(1.91)
Male	1.100	(1.14)	1.058	(0.55)	1.241***	(3.02)	0.980	(−0.23)
Immigrant Background	0.978	(−0.24)	1.122	(1.42)	0.946	(−0.70)	1.118	(1.45)
Firm in East Germany	1.546***	(2.78)	0.820	(−1.27)	1.668***	(3.85)	0.843	(−1.08)
Regional unemployment rate	1.005	(0.37)	1.018	(1.30)	0.982	(−1.42)	1.018	(1.31)
P: lower secondary schooling	0.648**	(−2.54)	0.951	(−0.26)	0.652***	(−2.76)	0.823	(−0.92)
P: intermediate schooling	0.658**	(−2.15)	1.133	(0.61)	0.643**	(−2.54)	0.995	(−0.02)
P: upper secondary schooling	0.597**	(−2.06)	1.047	(0.21)	0.596**	(−2.46)	0.918	(−0.35)
P: apprenticeship	0.957	(−0.49)	1.263**	(2.24)	0.977	(−0.30)	1.246**	(2.29)
P: master craftsman/technician	1.060	(0.39)	1.361**	(2.37)	0.994	(−0.05)	1.291**	(2.11)
P: higher education	1.307	(1.27)	1.458**	(2.27)	1.229	(1.22)	1.401**	(2.05)
P: employed	0.834*	(−1.68)	0.835	(−0.94)	0.895	(−1.29)	0.884	(−0.72)
P: self-employed	0.824	(−1.32)	0.907	(−0.48)	0.804*	(−1.75)	0.959	(−0.23)
Prose score	0.948	(−1.39)	1.095***	(2.81)	0.965	(−1.10)	1.088***	(2.70)
Document literacy score	0.997	(−0.07)	1.048	(1.34)	1.006	(0.20)	1.040	(1.18)
Numeracy score	1.031	(0.80)	0.985	(−0.45)	1.041	(1.28)	0.979	(−0.66)
High-cultural activity	0.818***	(−4.51)	1.102***	(2.64)	0.876***	(−3.57)	1.068*	(1.93)
Importance of work	0.936*	(−1.82)	0.951	(−0.78)	0.972	(−0.90)	0.972	(−0.53)
Importance of occupation	0.920**	(−2.39)	1.022	(0.48)	0.965	(−1.05)	1.021	(0.51)
Importance of friends	0.942	(−1.58)	1.039	(1.21)	0.981	(−0.64)	1.053*	(1.71)
Importance of family	0.969	(−0.74)	1.011	(0.29)	0.959	(−1.29)	0.995	(−0.15)
External locus of control	1.057	(1.48)	0.952	(−1.55)	1.019	(0.61)	0.953	(−1.60)
Internal locus of control	1.073	(1.45)	1.147***	(2.96)	1.033	(0.93)	1.123***	(2.88)

(table continued on following page)

	Model 3				Model 4			
	Non-participation		Frequency		Non-participation		Frequency	
Employment-related self-confidence	0.991	(−0.22)	1.034	(0.80)	0.998	(−0.06)	1.034	(0.87)
Lower secondary schooling	0.815	(−0.73)	1.173	(0.36)	0.826	(−0.72)	1.288	(0.56)
Intermediate schooling	0.506**	(−2.45)	0.946	(−0.14)	0.609*	(−1.91)	1.015	(0.04)
Upper secondary schooling	0.486**	(−2.53)	1.119	(0.27)	0.598*	(−1.92)	1.156	(0.35)
Apprenticeship	0.459***	(−5.50)	1.091	(0.37)	0.755*	(−1.90)	0.866	(−0.60)
Master craftsman/technician	0.238***	(−4.92)	0.920	(−0.33)	0.610**	(−2.38)	0.612*	(−1.86)
Higher education	0.179***	(−3.49)	1.214	(0.80)	0.864	(−0.79)	0.855	(−0.61)
Employment experience	0.974	(−1.18)	0.973**	(−2.06)	0.978	(−1.34)	0.977*	(−1.78)
Employment experience squared	1.000	(0.90)	1.000**	(2.18)	1.000*	(1.87)	1.000*	(1.91)
Civil service					0.728***	(−3.90)	1.114*	(1.73)
Working .5 full-time					0.735**	(−2.47)	1.587***	(2.90)
Working .75 full-time					0.566***	(−3.92)	1.889***	(3.84)
Working full time					0.487***	(−6.26)	1.723***	(3.69)
Induction required					0.818	(−1.56)	1.546**	(1.97)
Vocational training required					0.568***	(−4.99)	1.786***	(2.91)
Vocational schooling required					0.329***	(−6.01)	2.097***	(3.52)
Master craftsman/technician required					0.498***	(−4.04)	2.628***	(4.34)
Higher education required					0.213***	(−6.93)	2.449***	(4.10)
Firm size: 20-99					0.868*	(−1.93)	0.969	(−0.42)
Firm size: 100-199					0.958	(−0.45)	1.346	(1.42)
Firm size: 200-1,999					0.697***	(−4.52)	1.079	(0.92)
Firm size: 2,000 and more					0.597***	(−4.83)	1.091	(1.08)
Constant	4.410***	(4.37)	0.033***	(−6.94)	13.016***	(7.46)	0.015***	(−7.72)
$\ln(\alpha)$	2.532***	(9.12)			2.061***	(9.96)		
AIC	49303.7				48280.3			
BIC	49854.3				49032.6			
Wald-Statistic (χ^2)	200.283				285.084			
p-value	0.000				0.000			

Notes: ALWA, own calculations; 17254 observations; 6490 individuals; exponentiated coefficients; z-statistics in parentheses; cluster-robust standard errors; ***, **, * denote significance at 1%, 5%, 10%; additional controls: age-class, sex, immigrant background, East Germany and regional unemployment rate; reference category: parent of own sex without formal schooling or training and not employed, no own formal schooling or training, working ≤ 25% of full-time, no training required, firm size: 1-19 employees.

Finally, Model 4 also includes job and firm characteristics of the employment spell at hand. As expected they exhibit a strong influence on both the probability never to participate in on-the-job training and the number of training events per spell. Civil service employees are less likely to experience training abstinence and they participate in courses more often than employees in private firms. The higher the working time or the job requirement, the more likely and frequent is on-the-job training. Finally, on-the-job training is more likely in firms with more than 200 employees than in smaller firms. Given these results, it is not surprising that the results of Wald-tests depicted on Table 5 indicate joint significance of job characteristics. That corroborates the implications of theories of segmented labour markets. Although the job segment hypothesis is corroborated by these results in general, the firm size only influences training probability, but not its frequency. Being employed in a small firm might therefore be an obstacle to getting on-the-job training at all. As soon as a firm is able and determined to provide training at all, the number of training activities might no longer depend on its size.

Despite the strong explanatory power of the job and firm variables, most of the results from the more parsimonious models remain valid in the full specification given in Model 4, as can be seen in Table 5. Most importantly, on-the-job training is still strongly related to parental characteristics. This corroborates the intergenerational persistence hypothesis. Even in

Table 5: Wald tests of variable groups based on hypotheses and estimation results from Model 4

	Non-participation	p-value	Frequency	p-value
Family Background (H1)	13.819	0.087	24.110	0.002
Parental schooling	8.153	0.043	7.259	0.064
Parental training	2.175	0.537	6.560	0.087
Parental employment	3.100	0.212	1.298	0.522
Endowment (H2)	21.147	0.032	38.486	0.000
Cognition	3.077	0.380	8.795	0.032
Personality	17.963	0.022	24.300	0.002
Own education (H3)	24.688	0.000	14.065	0.029
Schooling	17.573	0.001	3.850	0.278
Training	7.082	0.069	10.088	0.018
Job characteristics (H4)	235.296	0.000	52.539	0.000
Working time	43.721	0.000	16.032	0.001
Job requirement	71.196	0.000	27.630	0.000
Firm size	36.129	0.000	3.664	0.453

Notes: ALWA, own calculations; the columns show χ^2-statistics and significance levels, respectively.

adulthood there is a channel through which an educated family background fosters educational attainment. A closer look at the coefficients reveals an even clearer picture of this influence. Whereas formal schooling achievements on behalf of the parents is associated with both a lower likelihood of on-the-job training abstinence and a higher frequency of training activities, parental training levels are only positively related to their offspring's frequency of on-the-job training. The latter might be due to networks of parents that influence decisions of employers, continued financial assistance or a higher cumulated wealth among the offspring of better educated and wealthier parents. The schooling of parents, on the other hand, might be better suited as an indicator to isolate wealth-related aspects from inherited cultural capital or intrinsic learning motivation. These different channels, however, cannot be investigated in more detail due to a lack of data.

Although not strongly significant, Wald-tests on cognition and personality generally support the endowment hypothesis. Endowments are less relevant for a lack of on-the-job training than for the frequency of training. While cognitive skills are significantly related to the frequency of training, they are not relevant for abstinence from training. Cognitive skills are therefore neither a prerequisite for receiving training in the first place, nor do they significantly foster participation when there is some other obstacle that leads to a lack of training. For those who receive training at all, cognitive skills do contribute to a higher number of training activities, which is in line with human capital theory and the endowment hypothesis. Personality traits are jointly significant for both equations.

Although the variables regarding the own formal education are jointly significant in Model 4, some coefficients no longer concur with the path dependency hypothesis. Both formal schooling and training make the absence of on-the-job training significantly less likely. Formal schooling,

however, no longer has a jointly significant influence on the frequency of on-the-job training. Existing formal training certificates even have a jointly significant negative relationship with the number of training activities per spell.

This counterintuitive result can be explained by the assumptions given in Section 2. If training participation depends on the financial means of a person, higher human capital measured by educational certificates leads to more training because liquidity constraints are less likely to be relevant. If, on the other hand, the wage is strongly related to the job segment one is employed in—as claimed by theory of segmented labour markets—job characteristics would be more important for remuneration than education. The training level as a proxy for the unobserved wage gets less important. If the job segment is controlled for, a complementary explanation gets relevant: A better educated and thus more productive worker has higher opportunity costs of training. This influences his investment decision towards less on-the-job training. Finally, human capital theory also offers an explanation of the negative relationship between formal training level and the number of on-the-job training activities per spell. A person that already has achieved a high level of human capital is more likely to refrain from additional investments. That is because she might already have achieved the optimal level of human capital, and investing more would yield less future returns than the actual costs.

6 Predictions and sensitivity analysis

Simulation of training participation by family background

To infer the economic relevance of the results, Table 6 shows predicted frequencies of on-the-job training per spell—denoted by y—by parental educational level as well as the probabilities of realising certain values for y. The predictions are computed based on the results from Model 4; the rows represent different parental educational levels. The column denoted by $E(y)$ represents the expected frequency of on-the-job training activities per spell, whereas the remaining columns show the probability to experience y training courses during a spell with y ranging from 0 to 3. Predictions for higher values of y are not shown due to their small share in the estimation sample. The confidence intervals in square brackets are estimated by bootstrapping with 1,000 replications.

The expected number of training activities per spell increases by parental education. This value is at least twice as high for people considered in rows three and four compared to people from a family background without any formal education in row one. This difference to the value of the offspring of uneducated backgrounds is significant for all other groups. However, the expected numbers of training activities per spell do not significantly differ between the three bottom groups as shown by the overlapping confidence intervals.

The same applies when considering the rest of the predicted values. Column 2 shows the predicted probability of experiencing no on-the-job training based on both equations of the

Table 6: Predicted frequency of training per spell (y) and probabilities of counts, respectively, by selected parental educational levels

	E(y) (1)	Pr(y=0) (2)	Pr(y=1) (3)	Pr(y=2) (4)	Pr(y=3) (5)
No schooling or training	0.727 [0.545-0.978]	0.753 [0.705-0.805]	0.095 [0.065-0.114]	0.052 [0.040-0.063]	0.032 [0.026-0.040]
Lower secondary schooling, apprenticeship	1.119 [1,046-1,352]	0.636 [0.606-0.648]	0.134 [0.120-0.141]	0.076 [0.072-0.081]	0.048 [0.046-0.052]
Intermediate schooling, master craftsman/technician	1.505 [1.169-2.045]	0.575 [0.522-0.619]	0.141 [0.110-0.159]	0.083 [0.072-0.093]	0.055 [0.050-0.062]
Upper secondary schooling, higher education	1.446 [1.129-1.948]	0.590 [0.537-0.633]	0.136 [0.107-0.154]	0.080 [0.069-0.091]	0.053 [0.047-0.060]

Notes: ALWA, own calculations; predictions based on Model 4; remaining covariates fixed at their mean; 95% confidence intervals in square brackets calculated by bootstrapping with 1000 replications.

zero-inflated negative binomial model. The result can be interpreted as the probability of experiencing no on-the-job training either because of a lack of opportunity or because of not using or having been offered on-the-job training although the opportunity would have existed. The predicted probability of a lack of on-the-job training for people from a family background without formal education is 75%. The same fate is less likely by 16 percentage points for people from, for instance, an academic family background. The conclusions when comparing the different parental educational levels are thus similar to those from column 1. The same applies for the remaining columns, albeit with smaller differences between the educational levels.

This shows that family background can play a significant role for on-the-job training, but this applies mainly to the lowest end of the parental educational distribution. A total lack of parental education is associated with low probabilities and frequencies of on-the-job training. Variation in parental education beyond that threshold is no major explanatory factor for non-formal training.

Sensitivity analysis

To check the robustness of the results, alternative specifications are tested. First of all, Model 4 is re-estimated using alternative specifications of family background. This is done to infer whether the results on this group of variables depend on assumptions regarding the measurement of the family background variables. The remaining control variables from Model 4 are included in the estimations but not shown in the tables hereafter. Depending on the information included in the respective models, the amount of missing values in the data and therefore the number of observations vary.

Table 7 shows the results of Model 4a which includes parental schooling and training as cumulated durations of educational activities instead of categorical educational levels. These durations are computed based on mean durations per educational level that have been tested

empirically for Germany (see e.g., Helberger, 1988). Other empirical applications can be found in Black et al. (2005) and Plug (2004).

Similar to the results from Model 4, more parental schooling is related to a smaller probability of non-participation and more parental training leads to a higher number of training events. Contrary to both the previous results and the intergenerational persistence hypothesis, the cumulated training duration is weakly associated with higher probability of non-participation in training. This could be interpreted as an inversely U-shaped relationship between parental training duration and non-participation in on-the-job training. Alternatively, it could be due to the fact that the variable representing parental cumulated training duration only has five different values. The influence of this variable thus can hardly be interpreted as linear. The Wald-test on the overall relevance of the family background, nevertheless, still corroborates the intergenerational persistence hypothesis.

Table 7: Determinants of non-participation in training and of the number of courses respectively, comparing different measures of family background, zero-inflated negative binomial regression, Probit inflation

	Model 4a				Model 4b			
	Non-participation		Frequency		Non-participation		Frequency	
P: schooling duration same sex	0.959***	(−2.60)	1.002	(0.11)				
P: training duration same sex	1.034*	(1.73)	1.041**	(2.19)				
P: employed	0.884	(−1.46)	0.902	(−0.62)				
P: self-employed	0.793*	(−1.91)	0.967	(−0.18)				
P: lower secondary					0.634**	(−2.44)	0.811	(−0.71)
P: intermediate schooling					0.639**	(−2.27)	0.982	(−0.06)
P: upper secondary schooling					0.701*	(−1.68)	0.926	(−0.25)
P: apprenticeship					0.867	(−1.55)	0.961	(−0.36)
P: master craftsman/technician					0.762**	(−2.36)	0.996	(−0.03)
P: higher education					0.821	(−1.42)	1.027	(0.18)
P: (self-)employed					0.929	(−0.18)	2.623***	(2.94)
Wald-statistic (χ^2)	246.897				298.573			
p-value	0.000				0.000			
Background (Wald-statistic)	23.737				39.737			
p-value	0.003				0.000			
Individuals	6490				6810			
Observations	17254				18050			

Notes: ALWA, own calculations; exponentiated coefficients; z-statistics in parentheses; cluster-robust standard errors; ***, **, * denote significance at 1%, 5%, 10%; omitted control variables similar to Model 4; Model 4a: parental schooling and training durations instead of categories; Model 4b: highest parental educational levels instead of that of the same sex as the respondent; reference categories: parent of own sex not employed (Model 4a), no parent formally educated or employed (Model 4b).

Whereas all parental characteristics so far have been based on information on the parent of the same sex as the respondent, Model 4b replicates Model 4 but uses the highest parental schooling and training levels given in the family as well as a dummy variable that indicates whether at least one parent had been employed when the respondent was 15 years old. Parental schooling and to a lesser extent also parental training make non-participation in training less likely, but none of these regressors are significantly associated with the frequency of on-the-job

training. Having had at least one parent that was employed or self-employed, on the other hand, strongly increases the number of training events. This should be due to larger financial means or more important networks of economically active parents. They are able to pass on wealth to their offspring which in turn faces less liquidity constraints for human capital investments. The role of proxy variables for parental wealth, which was occupied by parental training variables in Model 4, is thus transferred to the labour market status of the parents in Model 4b. All in all, the intergenerational persistence hypothesis can be sustained, and there is no indication on whether this specification should be preferred over the sex-role model specification in Model 4.

To test whether the influence of family background diminishes over time, spells from early in the respondents' life-courses are excluded. That way, own educational and labour market achievements gain relevance while the influence of the family background is diminished. Path dependency gets more important. This is tested by Model 4c in Table 8 in which all spells beginning before the age of 26 are excluded. Although the number of observations decreases markedly and less regressors regarding the family background remain significant taken by themselves, the Wald-test does not reject the joint significance of the family background variables.

A closer look reveals that this result is mainly driven by parental schooling rather than parental training. The influence of socialisation by the parents regarding the learning orientation seems to be more strongly and more persistently explained by parental schooling. It is also to be expected that the parental training level is more strongly related to unobserved wealth of the family. Since it is likely that financial support from parents decreases over the life-course of the offspring, the result from Model 4c also reflects a decreasing influence of unobserved wealth of the family on on-the-job training over time. The offspring gets more independent from its parents, even in financing his human capital investments.

Section 3 states that training decisions for civil service employees might follow rules that are not in line with human capital theory. Public employers are less inclined to decide on training provision based on cost-benefit considerations. Instead, their decisions are often oriented solely on administrative requirements of a given service grade. In that case, training participation would be determined more strongly by the training motivation of the given employee. To test this, Model 4d excludes employment spells in the civil service. The number of observations decreases again though less markedly than in Model 4c. The variables on parental training and labour market status are no longer significant but the Wald-test shows that the relevance of parental schooling is still significant.

These sensitivity analyses help us to refine the results and their interpretations but do not contradict the previous conclusions regarding the hypotheses. On-the-job training participation is related to both formal parental schooling and training. Whereas parental schooling is mainly associated with the general likelihood of training, parental training is particularly relevant for the number of training activities per spell. Both relationships diminish over the life-course of

Table 8: Determinants of non-participation in training and of the number of courses respectively, excluding spells starting below the age of 26 or of civil service respectively, zero-inflated negative binomial regression, Probit inflation

	Model 4c				Model 4d			
	Non-participation		Frequency		Non-participation		Frequency	
P: lower secondary schooling	0.588***	(−3.11)	0.786	(−1.04)	0.706**	(−2.12)	0.795	(−0.89)
P: intermediate schooling	0.752	(−1.50)	1.050	(0.20)	0.724*	(−1.80)	0.985	(−0.06)
P: upper secondary schooling	0.709	(−1.59)	0.983	(−0.07)	0.679*	(−1.83)	1.106	(0.36)
P: apprenticeship	0.992	(−0.08)	1.103	(1.11)	0.934	(−0.84)	1.108	(1.11)
P: master craftsman/technician	0.839	(−1.08)	1.114	(0.84)	0.913	(−0.73)	1.233	(1.59)
P: higher education	1.041	(0.23)	1.185	(1.01)	1.044	(0.27)	1.024	(0.16)
P: employed	1.070	(0.62)	1.084	(0.73)	1.003	(0.03)	1.178	(1.44)
P: self-employed	0.919	(−0.51)	1.169	(1.10)	0.896	(−0.82)	1.170	(1.08)
Wald-statistic (χ^2)	273.083				263.363			
p-value	0.000				0.000			
Background (Wald-statistic)	40.107				35.429			
p-value	0.001				0.003			
Individuals	4275				5391			
Observations	8942				13383			

Notes: ALWA, own calculations; exponentiated coefficients; z-statistics in parentheses; cluster-robust standard errors; ***, **, * denote significance at 1%, 5%, 10%; omitted control variables and reference categories similar to Model 4; Model 4c: exclusion of all spells beginning before the age of 26; Model 4d: exclusion of employment spells in the civil service.

the offspring.

7 Summary and conclusions

The aim of this study was to analyse whether people from low-qualified family backgrounds make up for their inherited lack of formal education by non-formal training. If they do not, one would have to conclude that a lack of intergenerational mobility in educational attainment is persistent over the life-course. Since education is highly relevant for both labour market and social participation this question is important for economic research as well as from a policy perspective.

Count data methods made analyses of the general probability of on-the-job training as well as its frequency possible. The negative-binomial regression model was able to deal with the substantive share of spells with a total lack of on-the-job training in the sample. All empirical analyses based on the newly available German ALWA-survey, a retrospective data set that includes a rich set of cross-sectional and longitudinal variables. The theoretical framework brought human capital theory and extensions thereof together with insights from genetics and psychology as well as with the theory of segmented labour markets.

Findings

The results reveal that learning inequality indeed is a lifelong phenomenon. In accordance with the intergenerational persistence hypothesis, family background is associated with a

person's human capital investment over the whole life. Growing up in a poorly educated household hampers one's prospects of formal and non-formal educational attainment even during adulthood.

However, the story is not as simple as that. Not only the education of parents influences someone's adult education. The endowment hypothesis stated that personality traits and cognitive skills, which are in part inheritable, can explain a substantive part of differences in on-the-job training participation between people from different family backgrounds. The analyses corroborate this hypothesis. One channel of intergenerational persistence therefore works through the inheritance of endowments that make educational attainment more likely.

The path dependency hypothesis stated that a higher level of formal education is associated with a higher likelihood and frequency of on-the-job training. This can only partly be sustained by the analyses. Existing formal schooling and training in fact make a total lack of on-the-job training less likely. They seem to be a prerequisite for further education. The frequency of on-the-job training per employment spell, though, is only influenced by previous formal training. When job and firm characteristics are controlled for, the existence of formal training is associated with a lower frequency of non-formal training. This might be due to high opportunity costs or to an already very high level of human capital. Additional non-formal training could exceed the optimal level of human capital, in which case such investments would not pay off in the long run.

Finally, the job segment hypothesis stated that employment in a favourable labour market segment leads to more on-the-job training. This is corroborated by the finding that both the likelihood and the frequency of training rise with working time and with job requirements. Firm size is only positively related to a higher likelihood of on-the-job training.

This study adds to the literature, for one thing, by showing that there are indeed factors that contribute to a lack of non-formal training—i.e. they constitute an obstacle to training. Other factors are associated with both the likelihood and frequency of training. The negative-binomial regression model has proven itself useful for more differentiated analyses than what would have been possible with the standard Poisson regression model.

For another thing and more importantly, the article brings together the question of intergenerational educational persistence with an analysis of the determinants of on-the-job training for the first time. By doing so, an important aspect of intergenerational mobility—or a lack thereof—could be analysed more comprehensively. The often cited blessings of lifelong learning are misleading as long as non-formal training does not contribute to a catching-up of people from low-qualified family backgrounds. They lack formal education themselves, and the gap in human capital even widens as they do not attain non-formal training often or intensively enough to close it.

Policy implications

These results allow several policy implications. They will be structured by the time-horizon in which remedies can take effect. In the short run human capital deficits could be reduced in part by incentives or financial support to participate in non-formal training limited to low-qualified people. Possible measures are, for instance, low interest loans for or tax deductibility of educational investments. There could also be incentives granted to firms that specifically invest in the human capital of their low-qualified employees.

These short-run remedies might be small steps towards more equal educational chances, but they certainly have a worse cost-benefit ratio than investments in human capital earlier in the life-course. Our results and findings reviewed, for instance, by Cunha and Heckman (2010) stress how important and profitable early investments in human capital can be. It is therefore not surprising that the following long-term policy implications resemble those from the literature on intergenerational persistence in formal eduction.

Family policy can play a major role in fostering intergenerational mobility. Possible measures include compulsory and cost-free kindergarten attendance or subsidised childcare facilities. Both would foster labour supply of parents—and expecially of mothers—of young children who would otherwise be unable to finance daycare for their children. The resulting increase in income would diminish financial constraints and allow additional investments in the human capital of the offspring. Moreover, well-conducted daycare facilities with well-educated staff offer favourable role-models and learning-oriented peer-groups. Both can foster the development of children's skills that are important future educational attainment and economic success (e.g., Heckman et al., 2010).

Educational policy has an equally important role to play. The comprehensive survey by Björklund and Salvanes (2010) identifies key obstacles in educational systems that lead to a lack of educational mobility. For one thing, the duration of compulsory school attendance proves to be negatively related to the influence of the family background on educational attainment. The longer children have to learn independently from parental decisions, the more human capital will be attained even after the period of compulsory school attendance. For another thing, the system of school tracking is important. The longer all pupils learn together before being separated into different tracks the lower the influence of the family background on early educational decisions.

Shortcomings and avenues for future research

Given the limitations of this analysis mentioned throughout the previous sections, this study constitutes merely a first step towards a better understanding of the long-term influence of family background on non-formal training participation. Further steps should include, for instance, attaining even better data on family background. The measure of labour market status used in this study is no sufficient index of the wealth parents were able and willing to

bequest to their offspring. A direct measure of wealth or more detailed information on labour market success of parents are therefore necessary.

Moreover, as the data set includes only cross-sectional measures of cognition and personality, the assumption of their time-invariability is crucial for the analysis. To overcome this shortcoming, repeated and more detailed measurements of these dimensions would be necessary. This analysis also hinges on the assumption that parental cognition and personality is inheritable to a certain degree. Although this has been ascertained by research in the fields of genetics and psychology, a direct test of this assumption in the context at hand would be preferable. This could be achieved by data that include measures of cognition and personality for both generations under consideration.

References

Anger, Silke and Guido Heineck (2010). 'Do smart parents raise smart children? The intergenerational transmission of cognitive abilities'. In: *Journal of Population Economics* 23.3, pp. 1105–1132.

Antoni, Manfred, Katrin Drasch, Corinna Kleinert, Britta Matthes, Michael Ruland, and Annette Trahms (2010). 'Arbeiten und Lernen im Wandel. Teil I: Überblick über die Studie'. In: *FDZ-Methodenreport* 05/2010. URL: http://doku.iab.de/fdz/reporte/2010/MR_05-10.pdf.

Arulampalam, Wiji and Alison L. Booth (2001). 'Learning and Earning: Do Multiple Training Events Pay? A Decade of Evidence from a Cohort of Young British Men'. In: *Economica* 68.271, pp. 379–400.

Asplund, Rita (2005). 'The Provision and Effects of Company Training: A Brief Review of the Literature'. In: *Nordic Journal of Political Economy* 31, pp. 47–73.

Backes-Gellner, Uschi, Johannes Mure, and Simone N. Tuor (2007). 'The Puzzle of Non-Participation in Continuing Training - An Empirical Study of Chronic vs. Temporary Non-Participation'. In: *Zeitschrift für Arbeitsmarktforschung* 40.2/3, pp. 295–311.

Bandura, Albert (1994). 'Self-Efficacy'. In: *Encyclopedia of Human Behavior*. Ed. by Vilayanur Subramanian Ramachandran. Vol. 4. San Diego et al.: Academic Press, pp. 71–81.

Becker, Gary S. (1962). 'Investment in Human Capital: a Theoretical Analysis'. In: *Journal of Political Economy* 70.S5, pp. 9–49.

Becker, Gary S. and Barry R. Chiswick (1966). 'Education and the Distribution of Earnings'. In: *American Economic Review* 56.1/2, pp. 358–369.

Becker, Gary S. and Nigel Tomes (1979). 'An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility'. In: *Journal of Political Economy* 87.6, pp. 1153–1189.

— (1986). 'Human Capital and the Rise and Fall of Families'. In: *Journal of Labor Economics* 4.3, S1–S39.

Björklund, Anders and Kjell G. Salvanes (2010). 'Education and Family Background: Mechanisms and Policies'. In: *Handbook of the Economics of Education*. Ed. by Eric A. Hanushek, Stephen J. Machin, and Ludger Wößmann. Vol. 3. Amsterdam et al.: North-Holland, forthcoming.

Black, Sandra E. and Paul J. Devereux (2010). 'Recent Developments in Intergenerational Mobility'. In: *NBER Working Paper* No. 15889. URL: http://www.nber.org/papers/w15889.pdf.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2005). 'Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital'. In: *American Economic Review* 95.1, pp. 437–449.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas Ter Weel (2008). 'The Economics and Psychology of Personality Traits'. In: *Journal of Human Resources* 43.4, pp. 972–1059.

Boudon, Raymond (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York et al.: John Wiley & Sons.

Bourdieu, Pierre (1986). 'The Forms of Capital'. In: *Handbook of Theory and Research for the Sociology of Education*. Ed. by John G. Richardson. New York et al.: Greenwood Press, pp. 241–258.

Bowles, Samuel and Herbert Gintis (2002). 'The Inheritance of Inequality'. In: *Journal of Economic Perspectives* 16.3, pp. 3–30.

Büchel, Felix and Markus Pannenberg (2004). 'Berufliche Weiterbildung in West- und Ostdeutschland. Teilnehmer, Struktur und individueller Ertrag'. In: *Zeitschrift für Arbeitsmarktforschung* 37.2, pp. 73–126.

Buchmann, Marlis, Markus König, Li Jiang Hong, and Stefan Sacchi (1999). *Weiterbildung und Beschäftigungschancen*. Chur et al.: Rüegger.

Cameron, Adrian C. and Pravin K. Trivedi (1986). 'Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests'. In: *Journal of Applied Econometrics* 1.1, pp. 29–53.

Cameron, Adrian Colin and Pravin K. Trivedi (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Cawley, John, James Heckman, and Edward Vytlacil (2001). 'Three observations on wages and measured cognitive ability'. In: *Labour Economics* 8.4, pp. 419–442.

Checchi, Daniele (2006). *The Economics of Education. Human Capital, Family Background and Inequality*. Cambridge et al.: Cambridge University Press.

Chevalier, Arnaud, Kevin Denny, and Dorren McMahon (2009). 'A Multi-country Study of Intergenerational Educational Mobility'. In: *Education and Inequality Across Europe*. Ed. by Peter Dolton, Rita Asplund, and Erling Barth. Cheltenham et al.: Edward Elgar. Chap. 12, pp. 260–281.

Colquitt, Jason A., Jeffrey A. LePine, and Raymond A. Noe (2000). 'Toward an Integrative Theory of Training Motivation: A Meta-Analytic Path Analysis of 20 Years of Research'. In: *Journal of Applied Psychology* 85.5, pp. 678–707.

Cunha, Flavio and James J. Heckman (2008). 'Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation'. In: *Journal of Human Resources* 43.4, pp. 738–782.

— (2010). *Investing in Our Young People*. NBER Working Paper 16201. URL: http://www.nber.org/papers/w16201.pdf.

Deary, Ian J., Steve Strand, Pauline Smith, and Cres Fernandes (2007). 'Intelligence and educational achievement'. In: *Intelligence* 35.1, pp. 13–21.

Dieckhoff, Martina (2007). 'Does it Work? The Effect of Continuing Training on Labour Market Outcomes: A Comparative Study of Germany, Denmark, and the United Kingdom'. In: *European Sociological Review* 23.3, pp. 295–308.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2010). 'Are Risk Aversion and Impatience Related to Cognitive Ability?' In: *American Economic Review* 100.3, pp. 1238–60.

Drasch, Katrin and Britta Matthes (2011). 'Improving retrospective life course data by combining modularized self-reports and event history calendars. Experiences from a large scale survey'. In: *Quality & Quantity* forthcoming.

European Commission (2000). *A Memorandum on Lifelong Learning*. SEC(2000) 1832. Brussels. URL: http://www.bologna-berlin2003.de/pdf/MemorandumEng.pdf.

Feinstein, Leon, Angela Lee Duckworth, and Ricardo Sabates (2004). *A Model of the Intergenerational Transmission of Educational Success*. Wider Benefits of Learning Research Report No. 10. URL: http://www.learningbenefits.net/Publications/ResReps/ResRep10.pdf.

Fouarge, Didier, Trudie Schils, and Andries de Grip (2010). *Why Do Low-Educated Workers Invest Less in Further Training?* IZA Discussion Paper No. 5180. URL: http://ftp.iza.org/dp5180.pdf.

Frederick, Shane (2005). 'Cognitive Reflection and Decision Making'. In: *Journal of Economic Perspectives* 19.4, pp. 25–42.

Frick, Joachim R., Markus M. Grabka, and Olaf Groh-Samberg (2007). *Economic Gains from Publicly Provided Education in Germany*. IZA Discussion Paper No. 2911. URL: http://ftp.iza.org/dp2911.pdf.

Gerner, Hans-Dieter and Jens Stegmaier (2009). 'Unsicherheit und betriebliche Weiterbildung. Eine empirische Analyse der Weiterbildungsaktivität unter Unsicherheit in KMU und Großbetrieben'. In: *Zeitschrift für Betriebswirtschaft* Special Issue 6/2009, pp. 135–163.

Green, David A. and W.C. Riddell (2003). 'Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation'. In: *Labour Economics* 10.2, pp. 165–184.

Grossman, Michael (2006). 'Education and nonmarket outcomes'. In: *Handbook of the Economics of Education*. Ed. by Eric A. Hanushek and Finis Welch. Vol. 1. Amsterdam et al.: North-Holland, pp. 577–633.

Hardin, James W. and Joseph M. Hilbe (2007). *Generalized Linear Models and Extensions*. 2nd ed. College Station: Stata Press.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006). 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior'. In: *Journal of Labor Economics* 23.3, pp. 411–482.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz (2010). 'The rate of return to the HighScope Perry Preschool Program'. In: *Journal of Public Economics* 94.1-2, pp. 114–128.

Heilbron, David C. (1994). 'Zero-Altered and Other Regression Models for Count Data with Added Zeros'. In: *Biometrical Journal* 36.5, pp. 531–547.

Heineck, Guido (2011). 'Does It Pay To Be Nice? Personality And Earnings In The UK'. In: *Industrial and Labor Relations Review* forthcoming.

Heineck, Guido and Silke Anger (2010). 'The returns to cognitive abilities and personality traits in Germany'. In: *Labour Economics* 17.3, pp. 535–546.

Heineck, Guido and Regina T. Riphahn (2009). 'Intergenerational Transmission of Educational Attainment in Germany - The Last Five Decades'. In: *Jahrbücher für Nationalökonomie und Statistik* 229.1, pp. 36–60.

Helberger, Christof (1988). 'Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie'. In: *Bildung, Beruf, Arbeitsmarkt*. Ed. by Horst Albach and Hans-Joachim Bodenhöfer. Berlin: Duncker & Humblot, pp. 151–170.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina (2007). 'The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends'. In: *B.E. Journal of Economic Analysis & Policy* 7.2, Article 10.

Hilbe, Joseph M. (2008). *Negative Binomial Regression*. Cambridge: Cambridge University Press.

Huber, Peter J. (1967). 'The behavior of maximum likelihood estimates under nonstandard conditions'. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Ed. by Lucien M. Le Cam and Jerzy Neyman. Vol. 1. Berkeley: University of California Press, pp. 221–233.

Jolliffe, Ian T. (2002). *Principal Component Analysis*. New York: Springer.

Kirby, Kris N., Gordon C. Winston, and Mariana Santiesteban (2005). 'Impatience and grades: Delay-discount rates correlate negatively with college GPA'. In: *Learning and Individual Differences* 15.3, pp. 213–222.

Lambert, Diane (1992). 'Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing'. In: *Technometrics* 34.1, pp. 1–14.

Leontaridi, Marianthi (1998). 'Segmented Labour Markets: Theory and Evidence'. In: *Journal of Economic Surveys* 12.1, pp. 63–101.

McCrae, Robert R. and Oliver P. John (1992). 'An Introduction to the Five-Factor Model and its Applications'. In: *Journal of Personality* 60.2, pp. 175–215.

Moulton, Brent R. (1990). 'An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units'. In: *Review of Economics and Statistics* 72.2, pp. 334–338.

Mullahy, John (1986). 'Specification and testing of some modified count data models'. In: *Journal of Econometrics* 33.3, pp. 341–365.

Pallas, Aaron M. (2004). 'Educational Transitions, Trajectories, and Pathways'. In: *Handbook of the Life Course*. Ed. by Jeylan T. Mortimer and Michael J. Shanahan. New York: Springer, pp. 165–184.

Pannenberg, Markus (1998). 'Weiterbildung, Betriebszugehörigkeit und Löhne: ökonomische Effekte des "timings" von Investitionen in die berufliche Weiterbildung'. In: *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*. Ed. by Friedhelm Pfeiffer and Winfried Pohlmeier. ZEW Wirtschaftsanalysen 31. Baden-Baden: Nomos, pp. 257–278.

— (2001). 'Schützt Weiterbildung on-the-job vor Arbeitslosigkeit?' In: *Bildung und Beschäftigung*. Ed. by Robert K. von Weizsäcker. Schriften des Vereins für Socialpolitik. Berlin: Duncker & Humblot, pp. 275–291.

Plomin, Robert, John C. DeFries, Gerald E. McClearn, and Peter McGuffin (2008). *Behavioral genetics*. 5th ed. New York: Worth.

Plug, Erik (2004). 'Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees'. In: *American Economic Review* 94.1, pp. 358–368.

Rotter, Julian B. (1966). 'Generalized expectancies for internal versus external control of reinforcement'. In: *Psychological Monographs* 80.1, pp. 1–28.

Schömann, Klaus and Janine Leschke (2008). 'Lebenslanges Lernen und soziale Inklusion - der Markt alleine wird's nicht richten'. In: *Bildung als Privileg - Erklärung und Befunde zu den Ursachen der Bildungsungleichheit*. Ed. by Rolf Becker and Wolfgang Lauterbach. 3rd ed. Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 347–383.

Schultz, Theodore W. (1961). 'Investment in Human Capital'. In: *American Economic Review* 51.2, pp. 1–17.

Sinnott, Jan D. (1994). 'Sex Roles'. In: *Encyclopedia of Human Behavior*. Ed. by Vilayanur Subramanian Ramachandran. Vol. 4. San Diego et al.: Academic Press, pp. 151–159.

Solon, Gary (2004). 'A model of intergenerational mobility variation over time and place'. In: *Generational income mobility in North America and Europe*. Ed. by Miles Corak. Cambridge: Cambridge University Press, pp. 38–47.

Taubman, Paul J. and Michael L. Wachter (1986). 'Segmented labor markets'. In: *Handbook of Labor Economics*. Ed. by Orley Ashenfelter. Vol. 2. Amsterdam et al., pp. 1183–1217.

Trostel, Philip, Ian Walker, and Paul Woolley (2002). 'Estimates of the economic return to schooling for 28 countries'. In: *Labour Economics* 9.1, pp. 1–16.

Vuong, Quang H. (1989). 'Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses'. In: *Econometrica* 57.2, pp. 307–333.

White, Halbert (1980). 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity'. In: *Econometrica* 48.4, pp. 817–838.

Winkelmann, Rainer (2008). *Econometric Analysis of Count Data*. 5th ed. Berlin: Springer.

Appendix

Table A.1: Sample statistics of independent variables

	Mean	s.d.	Min	Max
Age (in years)	28.00	(7.35)	18.00	51.33
Male (dummy)	0.50	(0.50)	0.00	1.00
Immigrant Background (dummy)	0.16	(0.37)	0.00	1.00
Prose literacy (score)	0.04	(1.00)	−3.85	2.14
Document literacy (score)	0.03	(0.99)	−4.00	2.28
Numeracy (score)	0.00	(0.98)	−3.23	2.39
High-cultural activity (score)	0.08	(0.96)	−2.00	3.67
Importance of work (score)	−0.05	(1.00)	−4.69	2.17
Importance of occupation (score)	0.12	(0.86)	−4.54	1.23
Importance of friends (score)	−0.08	(1.01)	−5.18	2.41
Importance of family (score)	0.05	(0.97)	−5.88	1.35
External locus of control (score)	−0.03	(0.98)	−4.40	2.37
Internal locus of control (score)	−0.00	(1.01)	−7.63	1.47
Employment related self-confidence (score)	0.04	(0.98)	−5.47	1.68
No schooling (dummy)	0.01	(0.08)	0.00	1.00
Lower secondary schooling (dummy)	0.21	(0.40)	0.00	1.00
Intermediate schooling (dummy)	0.40	(0.49)	0.00	1.00
Upper secondary schooling (dummy)	0.39	(0.49)	0.00	1.00
No training (dummy)	0.05	(0.21)	0.00	1.00
Apprenticeship (dummy)	0.66	(0.47)	0.00	1.00
Master craftsman/technician (dummy)	0.05	(0.22)	0.00	1.00

(table continued on following page)

	Mean	s.d.	Min	Max
Higher education (dummy)	0.24	(0.43)	0.00	1.00
Employment experience (in years)	4.85	(5.88)	0.00	33.00
P: no/unknown schooling degree (dummy)	0.03	(0.16)	0.00	1.00
P: lower secondary schooling (dummy)	0.64	(0.48)	0.00	1.00
P: intermediate schooling (dummy)	0.20	(0.40)	0.00	1.00
P: upper secondary schooling (dummy)	0.13	(0.34)	0.00	1.00
P: no/unknown vocational degree (dummy)	0.19	(0.39)	0.00	1.00
P: apprenticeship (dummy)	0.62	(0.49)	0.00	1.00
P: master craftsman/technician (dummy)	0.08	(0.28)	0.00	1.00
P: higher education (dummy)	0.11	(0.32)	0.00	1.00
P: not employed (dummy)	0.15	(0.36)	0.00	1.00
P: employed (dummy)	0.75	(0.44)	0.00	1.00
P: self-employed (dummy)	0.10	(0.30)	0.00	1.00
Firm in East Germany (dummy)	0.13	(0.33)	0.00	1.00
Regional unemployment rate	9.52	(3.58)	2.00	22.76
Civil service (dummy)	0.22	(0.41)	0.00	1.00
Working .25 full-time (dummy)	0.08	(0.28)	0.00	1.00
Working .5 full-time (dummy)	0.10	(0.30)	0.00	1.00
Working .75 full-time (dummy)	0.05	(0.23)	0.00	1.00
Working full time (dummy)	0.77	(0.42)	0.00	1.00
No training required (dummy)	0.11	(0.31)	0.00	1.00
Induction period required (dummy)	0.11	(0.32)	0.00	1.00
Vocational training required (dummy)	0.49	(0.50)	0.00	1.00
Vocational schooling required (dummy)	0.06	(0.24)	0.00	1.00
Master craftsman/technician required (dummy)	0.04	(0.19)	0.00	1.00
Higher education required (dummy)	0.18	(0.39)	0.00	1.00
Firm size: less than 5 (dummy)	0.08	(0.27)	0.00	1.00
Firm size: 5-9 (dummy)	0.13	(0.33)	0.00	1.00
Firm size: 10-19 (dummy)	0.13	(0.33)	0.00	1.00
Firm size: 20-99 (dummy)	0.23	(0.42)	0.00	1.00
Firm size: 100-199 (dummy)	0.11	(0.31)	0.00	1.00
Firm size: 200-1,999 (dummy)	0.21	(0.41)	0.00	1.00
Firm size: 2,000 and more (dummy)	0.12	(0.32)	0.00	1.00

Notes: ALWA, own unweighted calculations; 17254 observations.

Table A.2: Loss of observations due to exclusion restrictions and missing values based on Models 1 to 4 (not mutually exclusive)

	N of excluded spells	Share of total spells
Before exclusion restrictions	30,594	1.000
Spell from abroad	983	0.032
Former East Germany	2,033	0.066
Self-employed	2,034	0.066
Exclusion restrictions combined	5,046	0.165
Remaining after restrictions	25,707	1.000
Participation in on-the-job training	1,056	0.041
Training frequency per spell	1,150	0.045
Age class	1,072	0.042
Highest schooling degree	50	0.002
Highest training degree	672	0.026
Schooling of parent with same sex	1,912	0.074
Training of parent with same sex	1,532	0.060
Occupational status of parent with same sex	590	0.023
Prose score	601	0.023
Numeracy score	601	0.023
Document literacy score	601	0.023
High-cultural activity	49	0.002
Importance of work	457	0.018
Importance of occupation	457	0.018
Importance of friends	457	0.018
Importance of family	457	0.018
External locus of control	93	0.004
Internal locus of control	93	0.004
Employment related self-confidence	85	0.003
Civil service	577	0.022
Working time	280	0.011
Training requirement of job	243	0.009
Firm size	2,132	0.083
East Germany	654	0.025
Regional unemployment rate	928	0.036
All missings after restrictions	8,453	0.329
Remaining after all dropouts	17,254	0.564

Source: ALWA, own calculations