Four Essays on Human Capital

Von der Wirtschaftswissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften

- Doctor rerum politicarum -

genehmigte Dissertation

von

Diplom-Ökonom Philip Yang

geboren am 23.11.1981 in Chengdu

2014

Referent: Prof. Dr. Patrick A. Puhani

Korreferentin: Prof. Dr. Ulrike Grote

Tag der Promotion: 31.7.2014

Abstract

This dissertation deals with human capital. It includes four chapters, from whom the first three chapters are related to human capital in the context of workplace performance and the last is on human capital in the context of economics of education, particularly addressing human capital increasing effects of music activity. Chapter 1 evaluates the effects of age and tenure on careers and productivity in an internal labour market to find that the probability to make rewarded suggestions is inversely u-shaped related to age. Chapter 2 investigates the effects of age on the probability of workers participating in various employer-financed training measures. Also training participation is inversely u-shaped related with age. Chapter 3 studies the effect of training on employee promotions and workplace suggestions to find that training participation increases suggestions probability. Chapter 4 focuses on music as an alternative measure to increase human capital and studies the effects of musical activity on educational performance in childhood and adolescence. Music activity has a positive and long-term effect on educational attainment.

Keywords: productivity, age, suggestions, music, education

Zusammenfassung

Diese Dissertation beschäftigt sich mit Humankapital. Sie enthält vier Kapitel, wobei sich die ersten drei Kapitel mit Humankapital im Kontext von Arbeitsplätzen beschäftigen und das letzte Kapitel Humankapital im Kontext von Bildungsökonomik, insbesondere der Steigerung des Humankapitals durch die musikalische Aktivität behandelt. Kapitel 1 untersucht den Effekt des Alters auf Karrieren und Produktivität in einem internen Arbeitsmarkt und findet heraus, dass die Wahrscheinlichkeit prämierte Verbesserungsvorschläge zu machen mit dem Alter in einem umgekehrt u-förmigen Verhältnis steht. Kapitel 2 studiert den Effekt des Alters auf die Wahrscheinlichkeit an diversen, von dem Arbeitgeber bezahlten, Weiterbildungsmaßnahmen teilzunehmen. Auch die Wahrscheinlichkeit der Weiterbildungsteilnahme steht in einem umgekehrt u-förmigen Verhältnis zum Alter. Kapitel 3 behandelt den Effekt der Weiterbildung auf die Wahrscheinlichkeit einen Verbesserungsvorschlag zu machen und findet heraus, dass die Wahrscheinlichkeit einen Verbesserungsvorschlag zu machen mit der Teilnahme an einer Weiterbildung steigt. Kapitel 4 untersucht den Effekt der Musik, eine alternative Maßnahme zu Steigerung des Humankapitals, auf den Bildungserfolg in der Kindheit und in der Jugend. Musikalische Aktivität hat einen positiven und langfristigen Effekt auf den Bildungserfolg.

Keywords: Produktivität, Alter, Verbesserungsvorschlag, Musik, Bildung

Diese Arbeit widme ich meinen Eltern, Miri und Hanna

Acknowlegdement

Ich bedanke mich bei meinem Doktorvater, Patrick Puhani, für gute Betreuung, fachliche Unterstützung, konstruktive Kritik und kontinuierliche Hilfe. Durch ihn habe ich in den letzten Jahren das wissenschaftliche Arbeiten erst richtig kennengelernt. Gleichzeitig hat er mir ermöglicht, eigene Forschungsideen sowie Forschungsprojekte zu verwirklichen und stand mir dabei immer beratend zur Seite. Auch meiner Zweitgutachterin Ulrike Grote danke ich für ihre Unterstützung und Hilfsbereitschaft.

Ich danke der VW Stiftung für die finanzielle Unterstützung des Forschungsprojekts „Beschäftigungshemmnisse älterer Arbeitnehmer“. Im Rahmen dieses Kooperationsprojekts gilt mein besonderer Dank auch meinen Koautoren und Projektpartnern Christian Pfeifer, Simon Janssen und Uschi Backes-Gellner. Durch die Zusammenarbeit hatte ich die Gelegenheit, mich fachlich wie auch persönlich weiterzuentwickeln.

Ich bedanke mich bei allen Kollegen der Leibniz Universität Hannover. Vor allem Friederike, Sabine, Lausen, Timo, Karola, Fabian, Katja, Annika, Phan, Dirk, Anja, Pradyot, David, Atta, Kathrin, Sarah, Tabea, Antonia, Jörg, Reinhard, Alex und Konrad danke ich für die tolle gemeinsame Zeit.

Besonderer Dank gilt meinen Eltern für die guten Ratschläge und ihre kompromisslose Hilfsbereitschaft.

Schließlich danke ich Miri für all die Freude, Glück, Frieden und Liebe, die sie tagtäglich in mein Leben bringt.

Contents

List of Tables

List of Figures

Main introduction

This dissertation consists of four chapters. The first three chapters are related to an empirical research project that was supported by the VW Foundation as a part of the research initiative called "Employment problems of older employees". These three chapters cover various aspects of human capital in firms and are co-authored by Prof. Dr. Uschi Backes-Gellner, Prof. Dr. Christian Pfeifer, and Dr. Simon Janssen. Chapter 1 studies the effects of age and tenure on workers' careers and productivity in an internal labor market. Chapter 2 investigates the effects of age on the probability of workers participating in various employer-financed training measures. Chapter 3 studies the effect of training on employee promotions and workplace suggestions. Chapter 4 focuses on music as an alternative measure to increase human capital and studies the effects of musical activity on educational performance in childhood and adolescence.

Chapter 1 evaluates the effects of age and tenure on careers and productivity in an internal labor market. The data employed are administrative personnel records from a manufacturing firm in Germany; this firm has a distinct classification of wage groups and unique information on workers' productivity in the form of suggestions by workers. Using robust locally weighted regression, this study shows that the probability of promotion is highest for younger workers and for workers early in their careers; older workers are less likely to be promoted. The likelihood of making a rewarded suggestion is inversely u-shaped with respect to tenure and age. Furthermore, workers' abilities and firm-specific human capital determine their careers and productivity.

Chapter 2 analyzes the effect of age on training participation. A theoretical human capital model shows how amortization and screening affect age-related training probabilities. The results from theory are tested with a long panel dataset of male blue-collar workers to analyze the effects of age on the probability of participating in different employer-financed training measures. This study finds that the probability of participating in such training is inversely u-shaped with age and that longer training measures are undertaken earlier in both lives and careers. These findings are consistent with predictions from the human capital model that incorporates an amortization period and screening effects.

Chapter 3 evaluates the effects of employer-provided formal training on employee suggestions regarding how to improve productivity and on promotions among male blue-collar workers. Using more than 20 years of personnel data from four entry cohorts in a German company, this study makes use of various empirical methods to limit endogeneity caused by selection and reversed causality. The main finding is that workers are more likely to make rewarded suggestions and to be promoted after they have received formal training. However, the effect on suggestions is only short term. The probability of promotion is greatest directly after training but appears to be affected over the long term as well.

Chapter 4 analyzes the effect of music on cognition-related outcomes over the long term. Using multivariate regression under the assumption of conditional independence and individual fixed effects, this study shows that musical activity – either playing an instrument or singing – has a positive and significant effect on educational achievements during childhood and adolescence. In particular, music increases the predicted probability of being highly educated by around 20 percent. The music effect is highly significant for all specifications and music indicators, even holding individual and family characteristics constant. Alternative approaches, such as instrumental variables regressions, validate the robustness of the music coefficient. Music activity, particularly in early childhood, is important for human capital formation.

Chapter 1

Careers and productivity in an internal labor market*

Joint work with

Simon Janssen • Christian Pfeifer • Uschi Backes-Gellner

*This chapter is based on the publication Yang, P., Janssen, S., Pfeifer, and C., Backes-Gellner, U. (2013), Careers and productivity in an internal labor market, *Journal of Business Economics* 83, 121-143. Final publication is available under: http://link.springer.com/article/10.1007/s11573-012-0649-x

1.1 Introduction

The current demographic development is characterized by an aging population due to a steep drop in fertility rates and to increased longevity. As a consequence, relieving the potential economic burden of elderly ‘‘retirees’’ on the working population represents one of the most important challenges ahead. In addition to considerations regarding legislation related to retirement schemes, understanding workers’ disincentives and employer attitudes are of vital importance in improving the employability of the elderly (OECD 2006).

In this context, it is important to identify the degree to which career opportunities and individual productivity are dependent on workers’ age and tenure. Knowing the relationships among age, careers, and productivity allows for the development of adequate measures to reduce the barriers to employment for older workers. Despite the need for economic research on the relationship of age to careers and productivity, few studies have produced empirical evidence. Following criticism of the inability of early theoretical analyses to explain significant empirical findings on the organization of work within firms, the economic literature has increasingly focused on analyzing careers within organizations (Gibbons and Waldman 2006).

A large body of economic literature addresses wage and promotion dynamics in internal labor markets.[1] However, most studies address wage developments as the result of movements in hierarchical levels. Few studies have analyzed the age- related probability of promotion.[2] In general, the heterogeneity of the information on promotion represents a major challenge. Several authors have aggregated job levels through assorted job descriptions or the tasks of

[1] The seminal paper is written by Baker et al. (1994a, b). For reviews, see Gibbons and Waldman (2006), Gibbs (1995), Gibbons and Waldman (1999b).

[2] Dostie (2006) estimates the effects of age on wages and productivity. However, his results are limited by a relatively short observation period. Daveri and Maliranta (2006) use matched employer-employee data to observe the effects of age on wage and productivity.

workers to define promotions within hierarchies (Herpen et al. 2006). Also, structural stability demands no significant changes of job titles over time, and comparability among different studies demands a harmonized degree of simplicity (Baker et al. 1994a).

The relationship between age and individual productivity is mainly evaluated through the performance ratings[3] of employers, which are likely to be subjective. Alternatively, some studies use work samples, but the availability of data that truly represent the actual productivity of individuals[4] is scarce (Skirbekk 2003). A new branch of the literature utilizes matched employer-employee datasets.[5] Although these estimates are less likely to be subjective, a challenge in this approach is to identify age-related productivity effects that are not driven by other, unobserved characteristics of firms.

Few studies provide convincing data on individual or collective productivity. Jones (2010) and Weinberg and Galenson (2005) investigate publications in academic research. Other studies investigate performance in sports, such as the studies by Fair (2007) or Frank and Nüesch (2011). However, most of these studies observe very specialized branches of the economy, such as science, sports, or the arts, mainly because productivity measures are seldom available for workers in other, more common industries, such as manufacturing or service. One exception is Lazear (2000), who investigates the productivity of workers in a large auto glass company. Another notable exception is the paper by Börsch-Supan and Weiss (2007). These authors investigate the productivity of the work teams of the German car manufacturer Daimler AG by measuring the errors made by the teams during the production process.

---

[3] The seminal works using performance ratings are two studies by Medoff and Abraham (1980, 1981). A review of the relevant literature is written by Lin (2005).

[4] Oster and Hamermesh (1998) use the publications of researchers. Fair (1994) observes athletes. For a review, see Skirbekk (2003).

[5] See Abowd and Kramarz (1999) for a comprehensive review.

Therefore, our paper contributes to the existing literature by focusing on the effects of age and tenure on promotions and productivity. Furthermore, we incorporate information on workers' abilities to explain differences in the likelihood of promotion and performance. Our study differs in two main respects from existing studies of the effects of age and tenure on careers and productivity in internal labor markets. First, we use long, balanced panel data on workers within one firm and with a distinct classification of wage groups rather than aggregated job levels to derive information on promotions within hierarchies. Second, we proxy workers' productivity by using information from the internal suggestion scheme of the company, which is likely to be less subjective than survey data and which actually refers to individual performance to enhance the overall production process.

Based on a robust, locally weighted regression smoothing technique, our graphical analysis shows that promotion probability is initially high and decreases significantly with years of age and tenure. We provide an intuitive explanation for reduced promotion at the end of workers' careers by noting the lower productivity of older workers. Including ability in the analysis, which we proxy through workers' apprenticeship degrees, contributes further implications. First, workers with degrees have a higher promotion probability than workers without degrees, and workers of higher ability are more likely to make rewarded suggestions to enhance the production process. Second, workers who have received an internal apprenticeship have higher probabilities of promotion within hierarchies at the end of their careers compared to workers with external degrees. Furthermore, workers with internal degrees are much more likely to make suggestions than workers with external degrees.

This paper is structured as follows. The next section summarizes the previous empirical findings on the effects of employer-provided training. Section 3 describes the personnel dataset, provides descriptive statistics, and discusses the econometric framework. Section 4 presents the estimation results. The paper concludes with a short summary and a discussion of the results in Sect. 5.

1.2 Literature review

In recent decades, economic research has increasingly focused on understanding various aspects within firms, formerly treated as a ‘‘black box’’. Following criticism of the inability of existing theoretical analyses to explain significant empirical findings on the organization of work inside firms, the recent theoretical and empirical literature has made substantial progress in the field of internal labor markets. Along with new interpretations of internal labor market phenomena, the empirical work evaluates how existing theory matches findings, and theorists are developing new models to explain the patterns found in internal labor markets. Broad-based surveys of the now extensive literature are provided by Gibbons and Waldman (1999a, b), Lazear and Oyer (2006), Waldman (2007), and Lin (2005).

Two building blocks of economic theory, models of human capital acquisition and models of job assignments, provide rationales for the development of wages during careers. Wage growth with seniority is explained by four main concepts. First, wage growth with increasing seniority reflects firm-specific human capital acquisition, leading to higher productivity, which is shared between the firm and the worker (Becker 1962). Second, the efficiency wage concept of Becker and Stigler (1974) explains higher wages at the end of careers through employers’ interest in preventing shirking. Third, according to the deferred compensation theory proposed by Lazear (1979, 1981), wages are below the workers’ value of marginal products in the early career and above the workers’ value of marginal products at the end of the career because employers desire to retain workers with higher firm-specific capital. Fourth, Freeman (1977) explains wage growth over careers as a form of insurance. According to this explanation, workers' real wages will not fall until his actual ability becomes apparent in an attempt on the part of the employer to insure each worker against risk. Generally, experience in the labor market shapes wage growth, but workers' productivity is not dependent on seniority. Schooling, as a proxy for ability, determines workers' initial human capital when entering the firms. Workers with better education start at higher wages than workers with less schooling (Gibbons and Waldman 2006).

In theory, promotions serve two main purposes: as the assignment of workers to jobs within the firms and as incentives for effort and skill acquisition (Gibbons and Waldman 1999a). Job assignment models incorporate the concepts of learning and human capital acquisition. Learning refers to the employer's information about a worker's ability. Initially, firms know little about workers, but all employers gradually learn more about the ability of each worker. Whenever learning leads to the belief that a worker possesses high ability, workers will be promoted (Murphy 1986). Human capital acquisition in the context of job assignment predicts promotions whenever a worker has exceeded a certain effective ability threshold (Gibbons and Waldman 1999b). Workers' schooling, as a proxy for ability, serves to explain why certain workers accumulate human capital at a faster rate than others (Gibbons and Waldman 2006).

Because promotions are expected to be accompanied by large wage increases, firms are able to increase the effort and skill acquisition of workers by using promotions as incentives. Promotions increase the incentives for higher effort when higher productivity leads to subsequent promotions. Even when including the possibility of bribing supervisors, only those workers with higher output will be promoted (Fairburn and Malcomson 2003). Promotions increase skill acquisition if workers are more willing to invest in firm-specific human capital to move up in the hierarchy (Prendergast 1993). Alternatively, an up-or-out promotion rule forces workers to invest in firm-specific skills because non-promoted workers are subject to dismissal (Kahn and Huberman 1988).

There are many theoretical reasons to doubt that wage differentials result only from differences in productivity.[6] Therefore, seniority-related wage growth does not necessarily reflect increased productivity over the course of a career. Productivity increases at earlier stages of the career as workers accumulate human capital. After a certain threshold, individual productivity decreases with seniority and age due to health- or technology-related human capital depreciation (Mincer 1974).

The seminal empirical work on wages and promotions in an internal labor market consists of two in-depth analyses of white-collar firms in the US by Baker et al. (1994a, b). Since the publication of these analyses, several studies have followed, using white-collar workers (Lazear and Oyer 2006; Treble et al. 2001), blue-collar workers (Lazear 2000), and matched employer-employee data (McCue 1996; McLaughlin 1994). Three main results have been found with respect to promotions, wages, seniority, and workers' abilities.

First, real wages may decline with seniority, but nominal wages always increase with seniority. Workers who enter firms at a higher salary are more likely to receive significant real-wage increases with seniority. Those who receive the lowest entry salaries are likely to face a real wage decline. Promotions are associated with large wage increases, but wages also vary significantly within job levels (Lazear 1992; Baker et al. 1994a, b).

Second, promotion and wage growth are heavily dependent on a common underlying factor such as ability. A positive relationship between schooling, as a proxy for ability, and promotion probabilities has been found in several studies (Baker et al. 1994a, b; McCue

[6] E.g., incentive wages (Lazear 1979), efficiency wages (Akerlof and Yellen 1986), or sources of wage discrimination.

1996). Even after including controls such as experience and job levels, wage growth and promotion probability remain ability dependent (Habermalz 2006; Medoff and Abraham 1980, 1981).

Third, the process through which an employer receives more information about the ability of an employee (referred to as learning) plays an important role in determining wage growth and promotions. Firms use lower-level job performance to learn about the innate abilities of workers and thereby to facilitate promotion decisions. Promotions are less likely at higher job levels and at the beginning or the end of careers (Baker et al. 1994a, b; McCue 1996; Habermalz 2006). Firms prefer inside promotions over external recruitment when workers have similar abilities because firms know more about a worker with longer seniority. In addition, seniority-related promotion serves the purpose of rewarding higher corporate responsibility (Schottner and Thiele 2010; Chan 1996).

Empirical studies on the dependence of productivity on age or seniority find mixed evidence (Skirbekk 2003). Following the seminal work of Medoff and Abraham (1980, 1981), several studies use supervisors' job performance ratings to evaluate the effects of age and seniority on productivity, and these studies mostly find a lack of relationship or even negative associations between seniority and performance evaluation. Empirical work that uses work samples to identify productivity tends to find lower productivity levels among older workers.[7] Recently, several studies based on matched employer-employee data[8] have described an inverted u-shaped relationship between age and productivity. Work performance peaks when workers are between 30 and 40 years old. A substantial literature review on empirical studies based on matched employer-employee data has been written by Abowd and Kramarz (1999).

[7] See Skirbekk (2003) for a comprehensive survey.

[8] See Hellerstein et al. (1999), Hellerstein and Neumark (2004), Van Ours and Stoeldraijer (2010), and Goebel and Zwick (2009).

A recent paper by Börsch-Supan and Weiss (2007) investigates the work teams of a German car manufacturer to provide evidence regarding the relationship between the age composition of work teams and their team output. In contrast to former studies, these authors were able to rely on a very precise productivity measure in the form of the documented errors by work teams in the production process. In contrast to other papers, they do not find that productivity declines substantially with age.

1.3 The personnel dataset, variables, and methodology

We use the personnel data of a large German company from the energy sector, located in West Germany. The company is subject to a collective contract and has a works council. For reasons of data protection, we will neither name the company nor provide further company information. The data themselves contain a subsample of 438 blue-collar workers in the company's mining business, who entered the firm in four consecutive cohorts from 1976 to 1979. All of these workers stayed in the company throughout the entire observation period up to the year 2002. The balanced panel and extended length of the observation period reduce heterogeneity because we follow the workers throughout the majority of their work lives. The sample represents approximately 25 % of all employees in the company's operation unit and 3.5 % of the company's entire workforce. We observe information that has been collected through the monitoring system of the company. This approach offers tremendous advantages concerning the validity of the information. Whereas survey data from workers potentially suffer from measurement errors, as workers might under- or over-report information for various reasons, processed company data may be more reliable. A disadvantage of our balanced panel design is that we have no information about workers who left the firm.

We include only German male blue-collar workers without missing values in the variables used. As the only exception, we did not eliminate the information from 2002, from which the promotion variable is missing. The information on promotion is missing from that year because we do not observe the wage group of workers in the subsequent year, which would be necessary to generate the promotion variable. These restrictions reduce our sample by only 5 %. The final sample contains 10,544 observations of 415 different workers, who entered the company in four consecutive years from 1976 to 1979[9].

Lazear (2000) defines promotions in a hierarchy as movements into higher job levels[10] at a higher mean wage.[11] Despite this explicit definition, few empirical studies use an adequate measurement for hierarchies, mainly as a result of three problems. First, most studies have to aggregate job levels through numerous job descriptions or tasks of workers (Herpen et al. 2006). Therefore, structural stability demands that there be no significant changes in job titles and descriptions over time, and comparability among different studies requires a harmonized degree of simplicity when job titles are aggregated by job levels (Baker et al. 1994a). Second, changes in wages do not necessarily imply movements in hierarchical positions, limiting the use of information on wages to describe occupational hierarchy. Third, wages differ not only through dependence on hierarchical levels but also with respect to productivity and demographic factors, such as age, gender, education, and tenure (Grund 2002).

We benefit from the existence of a collective contract on wage groups in the company. The wage groups that we observe are comprehensively defined in the collective agreement. Higher wage groups correspond to higher levels in the hierarchy. In addition, the wage groups do not include significant differences of wages within one wage group. This factor

[9] 1976: Number of workers n = 105, panel length in years T = 27; 1977 n = 96, T = 26; 1978 n = 77, T = 25; 1979 n = 137, T = 24.

[10] Baker et al. (1994a) identify levels through job descriptions. Herpen et al. (2006) use job ratings.

[11] Overall, empirical studies on internal labor markets find significant positive effects of promotions on pay, with differences with respect to convexity and size (Grund 2002).

eliminates large within-group variations of pay due to unobservable factors. Our wage group variable ranges from 0 to 19, where 0 reflects the lowest wage group and 19 the highest. Wage group 0 represents a special case because all workers under the age of 18 without apprenticeship degrees are classified in this group until they reach maturity. In the analysis of age- and tenure-related wage group probability, we include this group because they do represent workers with the lowest pay and lowest hierarchy.

We use promotions within a hierarchy to describe careers within the organization. Promotions are important because firms are able to move workers of higher ability into higher levels (Fairburn and Malcomson 2003). In addition, promotions serve as an incentive mechanism for workers to increase productivity and human capital because an upward movement in the hierarchy ensures significantly higher pay (Prendergast 1993). Our analysis of promotion probability serves two main functions. First, it serves as a robustness check for the analysis of wage groups. Second, it provides further insights because we are able to identify age- and tenure- related differences in promotion probabilities, which are not observable solely through the analysis on wage groups. The promotion variable is binary and takes the value of one if a worker is in a higher wage group in the following year.

By definition, this excludes observations in the year 2002 for promotions because we do not have information on the following year. Additionally, we do not consider a shift from wage group 0 to a higher wage group as a promotion in hierarchy. Wage group 0 represents immature workers without apprenticeship degrees. After 18 years of age, these workers automatically shift into higher wage groups. We do not consider this shift as a promotion in hierarchy because it is specifically age- related in comparison to the other movements. This approach complements the existing literature on promotions in hierarchy because our unique promotion variable factors in both higher pay and a higher job level. The changes in job title and pay occur simultaneously (compare Baker et al. 1994a, b).

As a proxy for productivity, we use information on workers' contributions of suggestions to the company. Suggestions refer to the workers' effort to enhance the production process of the company and to increase the overall performance of the workforce. These ideas reflect innovative work behavior and the ability of workers to improve the quantity and quality of production for the company, rewarded with some type of compensation (Frese et al. 1999; Ford 1996). Our proxy for individual productivity reflects the value added for the firm in that suggestions serve to enhance the production process and to lower the production costs. Workers are monetarily paid in proportion to the money saved. Therefore, we only include suggestions leading to a bonus for the worker who made the suggestion.

Our suggestion variable is binary and is equal to one if a worker has made a suggestion to the firm that has led to a monetary reward for the worker. Unfortunately, we do not know the exact reward for each suggestion. However, the compensation in the form of a monetary reward indicates a certain quality of the propositions made. These unique data on suggestions differentiate our paper from previous empirical studies, which mainly use wages (Skirbekk, 2003), aggregated matched employer-employee data (Hellerstein et al. 1999; Van Ours and Stoeldraijer 2010; Goebel and Zwick 2009), performance evaluations (see Medoff and Abraham 1980, 1981), or actual manufactures (Lazear 2000) to analyze tenure- or age-related productivity.

We investigate careers and productivity with respect to years of tenure, age, and ability. Because the length of observation is relatively long, we follow workers over the majority of their work lives. The balanced panel design provides information on all workers and the changes over time for each worker, which limits the influence of other, unobserved characteristics. Because not all workers are of the same age when entering the firm, we use both continuous tenure and age information as explanatory variables.[12]

As a proxy for ability, we include information on apprenticeship degrees. The data allow us to distinguish between workers with no degrees, with external degrees, and with internal apprenticeships. We could simply control for apprenticeship degree through a binary apprenticeship degree variable and not distinguish between internal and external degrees if external and internal apprenticeship degree had identical implications for workers' abilities. However, in that case, we would fail to account for differences in firm-specific human capital and worker-firm attachment.

Internal candidates are expected to be favored for promotions and human capital investments due to a higher level of firm-specific human capital and as a reward for a higher level of corporate identity (Schöttner and Thiele 2010; Baker et al. 1994a, b). Consequently, we have created three dummy variables reflecting the apprenticeship groups. Table 1.1 presents the summary statistics for the variables of interest.

[12] The use of continuous age and tenure information serves the purpose of detailed descriptive analysis.

Table 1.1. Descriptive statistics

	Mean	SD	Min, Max	Cases
Explained variables				
Wage group	6.730	2.828	0, 19	10,544
Promotion (dummy)	0.083	0.276	0, 1	843
Suggestion (dummy)	0.035	0.183	0, 1	367
Explanatory variables				
Age	31.930	7.854	15, 54	10,544
Tenure	13.231	7.377	1, 27	10,544
Apprenticeship degree inside firm (dummy)	0.251	0.434	0, 1	2,651
Apprenticeship degree outside firm (dummy)	0.480	0.500	0, 1	5,064
Apprenticeship (reference: no degree)				2,829

Number of observation is 10,544 from 415 workers in a balanced panel design. The only exception constitutes the promotion variable, which is observed for all years except 2002. The number of observations for the promotion variable is 10,129

The average wage group is 6.73, with a standard deviation of 2.83, which implies that the majority of workers are settled in the lower segment of wage groups. Nevertheless, movements along wage groups are not uncommon, given a relatively large variability. The mean of promotion indicates that 8.3 % of workers are promoted yearly, which corresponds to 843 promotion cases over the entire observation period. Compared to other empirical studies, promotions may seem rare for our sample. However, it must be noted that we observe blue-collar workers, for whom promotions are less likely than in the white-collar sector. On average, 3.4 % of all workers make a suggestion that is monetarily rewarded, for a total of 367 suggestions. The lowest age observation in our sample is 15 years, and the oldest is 54 years. The mean age of the entire sample is 31.93 years, with a standard deviation of 7.85 years. We follow workers from the first year of tenure up to observations of workers who have spent 27 years in the company. On average, this range denotes 13.23 years of tenure, with a standard deviation of 7.28 years. Every fourth worker has completed an apprenticeship in the company. Approximately half of the workers hold an external apprenticeship degree, and 25 % of the workforce in the sample has never completed an apprenticeship.

To investigate careers in an internal labor market, we rely on non-parametric estimation methods. From a theoretical perspective, we cannot assume that the relation between tenure or age and the workers' productivity is linear. For example, Mincer (1974) assumes that workers invest substantially in their human capital in the early stages of their careers but reduce their investments (linearly) as the point of retirement comes closer. Consequently, the workers' productivity is a function of their age or tenure that increases at a decreasing rate. Additionally, skills might depreciate, either because the aging worker might lose motor or cognitive skills or because new technologies render obsolete the skills associated with outdated technologies (Grip and Loo 2002). As a result, the productivity of older workers might decrease, and older or long- tenured workers might even become less productive than their younger counterparts. Therefore, estimating the relationship between tenure and our dependent variables of wage categories, promotions and suggestions in a linear framework ignores such influences and might lead to poorly specified empirical models. One potential remedy would be to assume a particular functional form, such as, for example, a quadratic polynomial. However, because no straightforward theoretical model could guide us in choosing the right empirical model specification, we decided to let the data speak and to rely on a non-parametric estimation approach.

Obviously, the application of non-parametric estimators comes at a cost. We cannot control for unobserved heterogeneity, and the curse of dimensionality makes it difficult to add further control variables. However, because the nature of our study is descriptive and explorative, these drawbacks are less significant than the advantages of a non-parametric approach.

In our particular case, we carry out the locally weighted regression. This smoothing procedure was introduced by Cleveland (1979, 1981). Applications are described in Cox (2005), Royston (1991), and Royston and Cox (2005). Locally weighted regressions represent a method of smoothing scatterplots. The smoothing method is nonparametric in the sense that no a priori specification of the relationship between explanatory and explained variables is required (Jacoby 2000). Basically, we create new smoothed dependent variables by running regressions for each set of dependent and independent variables. Thus, the regressions are weighted so that the central regression pair receives the highest weight. Based on the distance to the central regression pair, other observations receive less weight with increasing distance to the central regression pair. The estimated regression line is used to predict the smoothed value for the dependent variable. This procedure is performed for all smoothed values, corresponding to as many separate weighted regressions as there are points in the data.

Specifically, we consider a dependent variable y_i, either the wage group, promotion, or suggestion, to be dependent on an explanatory variable x_i, either years of age or years of tenure, with $x_i \prec x_{i+1}$ for i = 1, . . . ,N – 1. For every dependent variable y_i, a weighted regression predicts a smoothed value $\hat{y}_i$ at each x_i. In other words, in contrast to a classical OLS regression, we do not consider all values of x_j (whereby j $j \in [1; N]$ and $j \neq i$) with equal weights to predict $\hat{y}_i$. Rather, we select a subset of observations and give x_j a lower weight if the distance between x_j and x_i is large and a greater weight if the distance between x_j and x_i is small. The subset of observations used to predict each $\hat{y}_i$ lies between $i_- = \max(1, i - z)$ and $i_+ = \min(i + Z, N)$ with $z = [(N * bwith - 0.5)/2]$, where *bwith* is a bandwidth that we have to specify. We use a bandwidth of 0.8, which means that 80 % of the data are used for each regression, reducing the influence of the observations of the extreme left on the fitted values on the extreme right.

This procedure is carried out for every x_i in our data, such that we estimate N different weighted regressions of the following form:

$$\sum_{k=1}^{N} w_k(x_i)(y_k - \beta_0 - \beta_1 x_k)^2, \tag{1}$$

with k = 1,…, *N*. In our particular case, the weighting function is of the following form.

$$w_j = \{1 - (\frac{|x_j - x_i|}{\Delta})^3\}^3, \tag{2}$$

where $\Delta = 1.0001\max(x_{i+} — xi, xi — x_{i—})$.

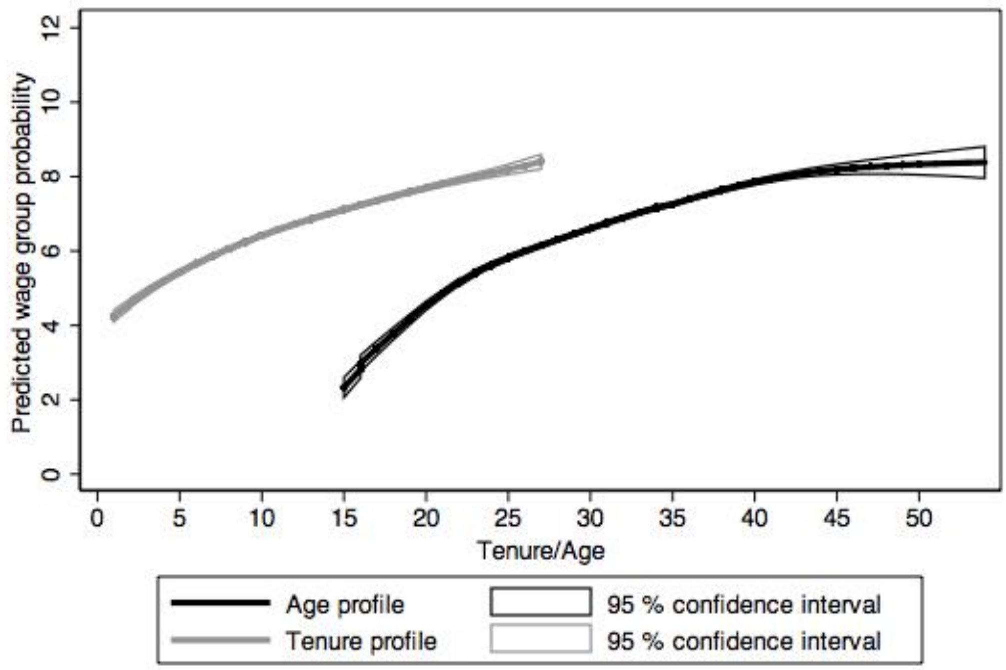
To estimate the confidence bands, we follow Cleveland (1979). In more detail, assume that $r_k(x_i)$ is a function that satisfies $\hat{y}_i = \sum_{k=1}^{N} r_k(x_i) z_k$ for every predicted value of y_i. Moreover, the (i,k)th element of the matrix R is $r_k(x_i)$. Following Cleveland (1979), the covariance matrix is now $\sigma^2 RR$, and we can predict the standard errors according to $\hat{\sigma}^2(\sum_{k=1}^{N} r^2 \mathrm{k}(x_i))^{1/2}$. The estimate of σ is $\hat{\sigma} = (t^{-1} \sum_{K=1}^{N} \varepsilon^2)^{1/2}$ with $\varepsilon^2 = y_i - \hat{y}_i$ and $t = tr((I — R)(I — R)')$.

1.4 Careers within the firm

We initiate our descriptive analysis on careers by estimating weighted regressions at each observation point for wage groups with respect to years of tenure and years of age. As we mentioned above, ap plying a non-parametric regression comes at the cost that we cannot incorporate additional control variables, nor can we apply panel data regressions. Therefore, as a robustness check, we estimated random-effects models to control for the unobserved heterogeneity of workers, which includes additional control variables. The results are presented in the appendix. Our main objective is to identify how the likelihood of being in a certain wage group relates to workers' age or tenure. We are also interested in developing insights regarding entry positions within the company. We predict wage group probabilities based on both age and tenure because not all workers entered the firm at the same age. Figure 1.1 plots the results.

The wage group probability is positively associated with both years of tenure and years of age. We observe considerable career opportunities at the beginning of work life, which diminish with increasing years of age and years of tenure. Both curves are steeper in the early career and flatten in the later stages of careers. After 40 years of age, the wage group-age relationship is almost inelastic. However, the confidence bands become broader for the estimates beyond the age of 40, indicating that these values are less precisely estimated. Obviously, we have fewer observations for this age group. The diminishing correlation is less stringent for tenure-related wage group probability. However, the effect of tenure on wage groups decreases after approximately 10 years of tenure.

Generally, entry positions, depending on age and tenure, are between wage group 2 and wage group 5, which represent the lower bound of the wage groups. Predicted entry wage groups are higher in the initial year of tenure than for the lowest age of workers. This difference may stem from two factors. First, immature workers without apprenticeships are located in wage group 0 until they have turned 18, and they lower the predicted entry wage group probability. Second, external candidates may have entered the firm in significantly higher positions in the hierarchy at older ages, leading to higher entry wage groups with respect to age.

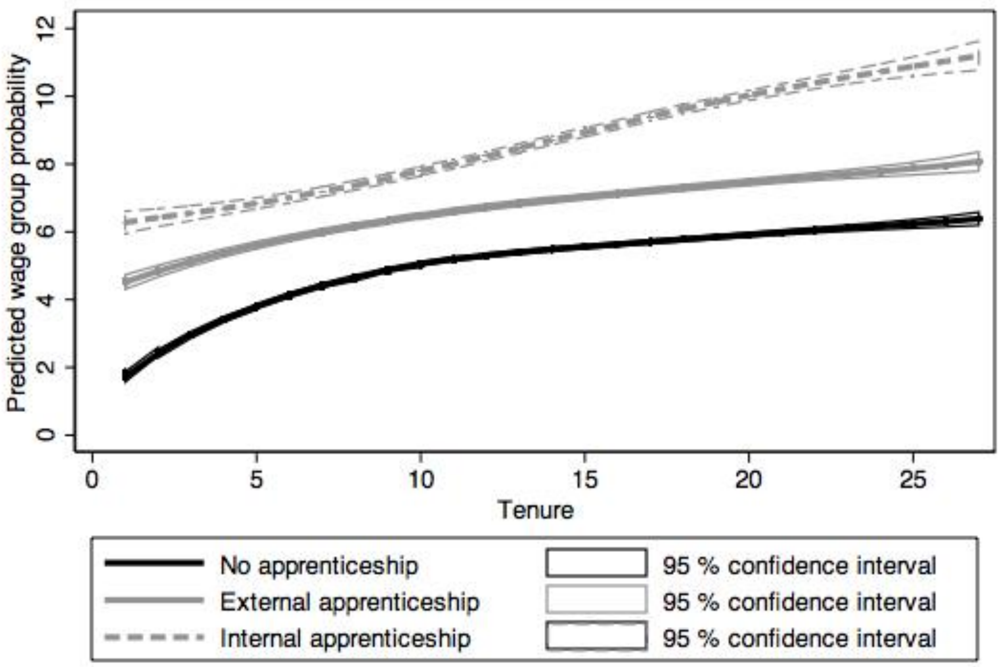


Note: Own calculations. The solid (no represents the predicted values from a locally weighted regression with bandwith 0.8.

Figure 1.1. Wage groups and tenure/age from locally weighted regressions

To pursue these ideas, we estimated separate locally weighted regressions for tenure-related wage group probability depending on ability. We use apprenticeship degrees to proxy ability and distinguish among workers with no apprenticeship degrees, workers with external apprenticeship degrees, and workers with internal apprenticeship degrees. Figure 1.2 plots the estimated results from robust weighted regressions at each observation point for these groups.

We observe substantial differences among the apprenticeship categories. The predicted wage groups for workers with internal apprenticeships, also referred to as internals, are significantly higher than the predicted wage groups for external candidates throughout the career. Workers with external apprenticeship degrees, also referred to as externals, begin at significantly higher wage groups than workers without apprenticeships. The curves do not intersect, indicating higher wage group probabilities throughout the work lives of members of the former group. In the first 10 years of tenure, the curves for internals and externals have almost identical slopes, whereas the predicted wage group differences between externals and workers without apprenticeships are large in the first years and then slowly diminish. Thereafter, the difference in wage group probabilities between workers with internals and externals increases, and the difference between workers with external degrees and no apprenticeship degrees remains constant.



Note: Own calculations. The solid line represents the predicted values from a locally weighted regression with bandwith 0.8.

Figure 1.2. Wage groups and tenure–different profiles from locally weighted regressions

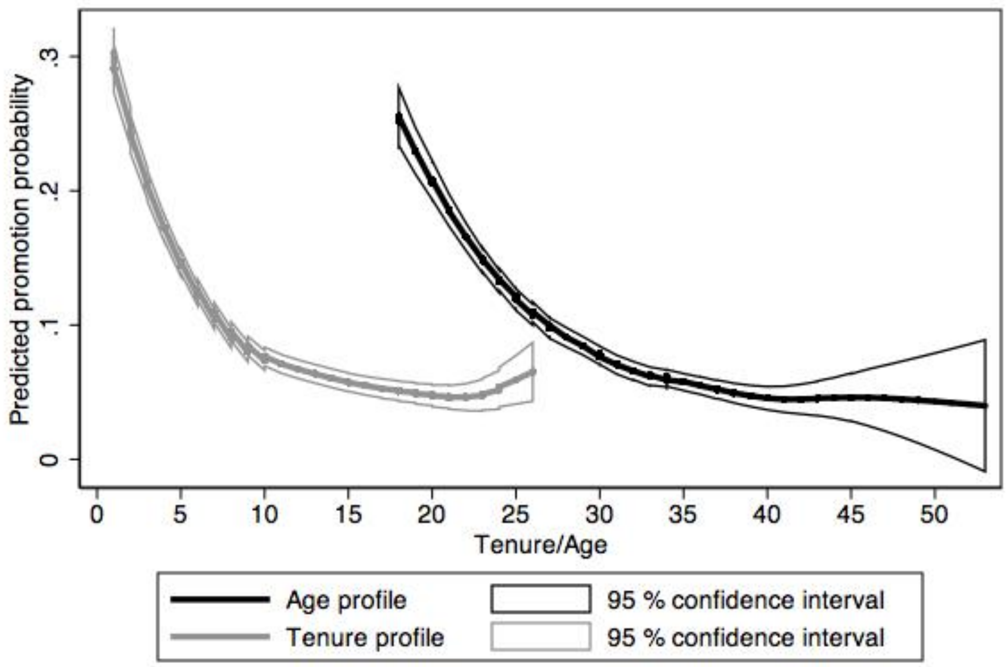
These findings imply that workers with apprenticeship degrees, which we consider as a proxy for higher ability, are much more likely to be in higher wage groups than workers without apprenticeship degrees. One explanation is that workers who have higher abilities are more productive. Alternatively, ability may determine the amount of human capital of workers and the speed and capability of workers to accumulate firm-specific human capital. One potential critique of our human capital- or ability-related interpretation of the results is that some wage groups in the collective agreement are a function of the worker's qualification. Given such hierarchical requirements, workers with an apprenticeship degree might not necessarily have more human capital or be more able than workers without an apprenticeship degree but might only be on a higher hierarchical level because of the formal definition of the wage group. However, the hierarchical qualification definitions provide only suggestions for the initial entry position of newly hired workers. The hierarchical qualification definitions are not binding, and workers can freely move between hierarchical levels without obtaining the required qualification. Therefore, we observe a substantial number of workers without apprenticeship degrees in hierarchical levels with an initial apprenticeship requirement. Consequently, we argue in favor of a human capital- or ability-related explanation for the observed career pattern, in particular, because even long-tenured workers with an apprenticeship degree are more likely to be on higher hierarchical levels than long tenured workers without an apprenticeship degree.

When comparing workers with apprenticeship degrees, workers with internal degrees are much more likely to be in higher wage groups than workers with external degrees. Incumbents with an internal apprenticeship degree are likely to be favored for promotions because they represent more firm-specific human capital. Another explanation is that the firm has learned more about internal workers in comparison to workers with external degrees or that the firm rewards higher corporate responsibility.

Following the descriptive results from the wage groups, we estimate locally weighted regressions for promotions. The analysis of age- and tenure-related promotion probabilities serves as a robustness check for the results on wage groups and provides further insights into actual promotion dynamics within the firm because it also denotes the slope of the wage group curves. Figure 1.3 displays promotion probabilities with respect to years of age and years of tenure.

The findings on the predicted promotion probabilities with respect to age and tenure support the previous results on age- and tenure-related wage groups. Promotions within hierarchies are most likely during the early career and diminish significantly with increasing years of age and years of tenure. For both factors, tenure and age, promotions in hierarchies are most likely when workers enter the firm. Afterward, promotion probability decreases significantly up to a level at which the relationship between promotion and age or tenure is almost inelastic until the end of workers' careers. After 10 years of tenure, promotion and seniority remain The findings on the predicted promotion probabilities with respect to age and tenure support the previous results on age- and tenure-related wage groups. Promotions within hierarchies are most likely during the early career and diminish significantly with increasing years of age and years of tenure. For both factors, tenure and age, promotions in hierarchies are most likely when workers enter the firm. Afterward, promotion probability decreases significantly up to a level at which the relationship between promotion and age or tenure is almost inelastic until the end of workers' careers. After 10 years of tenure, promotion and seniority remain almost uncorrelated at the lowest probability for promotion throughout the rest of the work life. Workers from the age of 40 years upward face stagnant, minor promotion chances. Again, the promotion probabilities are less precisely estimated for older age cohorts and for workers with longer tenure.

As shown above, the likelihood of promotion is higher in the initial years of tenure than in the initial years of age. Because we adjusted the promotion variable to exclude strictly age-related promotions, such as the shift out of wage group 0 at maturity, the main difference might lie in the heterogeneous promotion probabilities between workers with external degrees and workers with internal degrees. Promotions may be more likely in the first years of tenure than for the youngest workers if the workers with external degrees have higher initial promotion probabilities than workers with internal degrees.



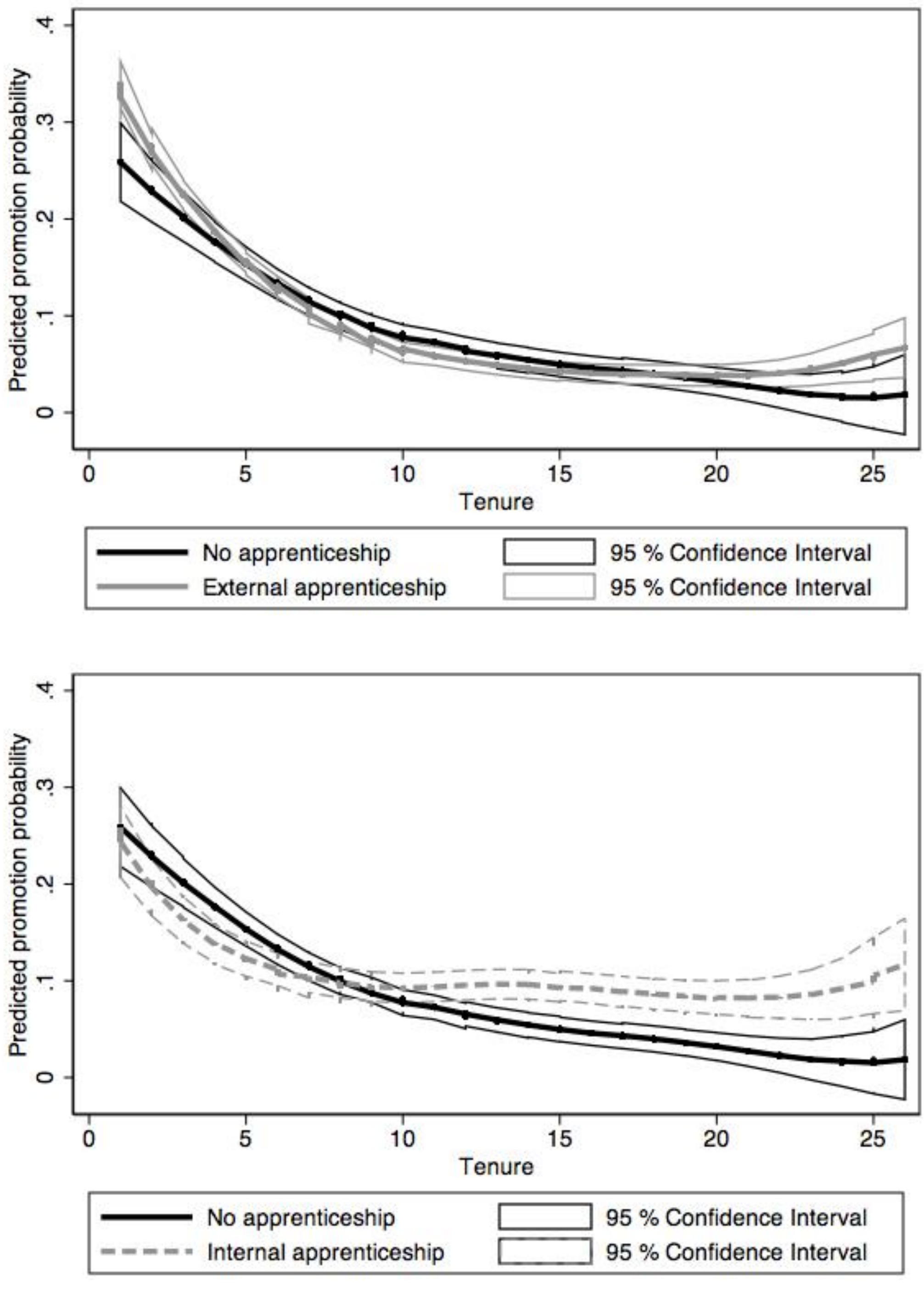
Note: Own calculations. The solid line represents the predicted values from a locally weighted regression with bandwith 0.8.

Figure 1.3. Promotion and tenure/age from locally weighted regressions

We estimate promotion probabilities for years of tenure for different apprenticeship degrees to evaluate distinctions with respect to ability and to test the previous assumption. Again, we differentiate among workers without apprenticeship degrees, with external apprenticeship degrees, and with internal apprenticeship degrees. This differentiation provides insight into the effect of ability on the likelihood of promotions within hierarchies. Moreover, we observe how promotions differ for workers with external degrees in comparison to workers with internal degrees. Figure 1.4a compares the estimated results from weighted regressions for individuals without an apprenticeship degree to the results for individuals with an external apprenticeship degree.

We observe substantial differences in the tenure-promotion relationship when differentiating among workers with respect to apprenticeship degrees. Initially, workers with external apprenticeship degrees have a slightly higher promotion probability than workers without an apprenticeship degree. However, the confidence bands of both estimations overlap so that we cannot expect the differences to be significant. Afterward, the predicted curve for apprentices without an apprenticeship degree and the curve for apprentices with an external apprenticeship degree become more similar, and the confidence bands remain overlapping. In contrast to Fig. 1.4a, 1.4b compares workers who underwent their apprenticeship training within the firm with workers who did not obtain an apprenticeship degree at all.

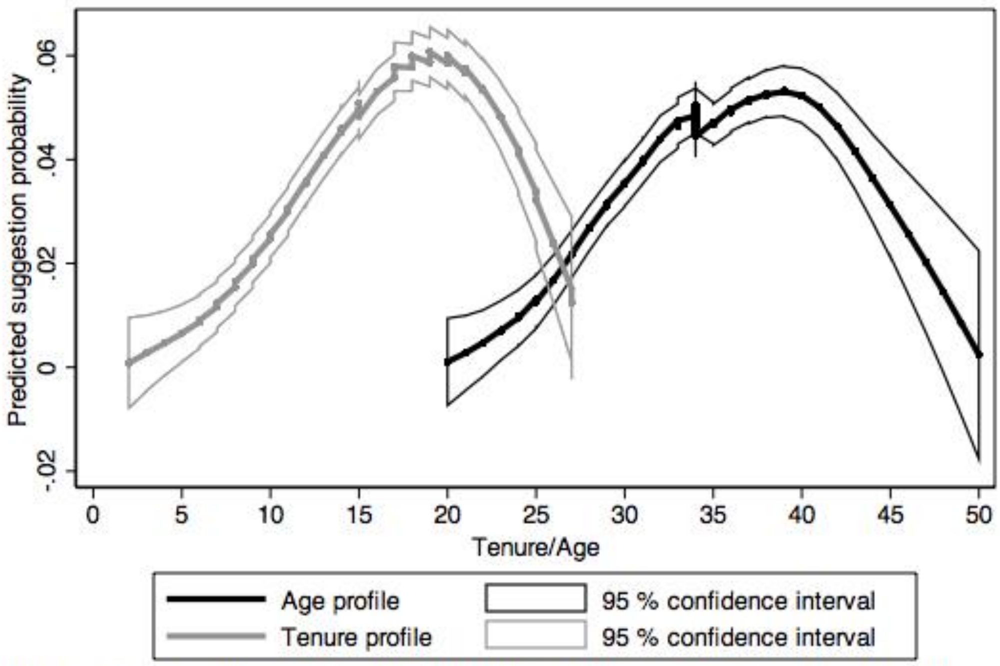
At early career stages, the promotion probability of workers who obtained their apprenticeship degree within the firm does not differ substantially from workers who did not obtain an apprenticeship degree. Specifically, the confidence intervals of the estimated promotion probability overlap in the first years of tenure. In contrast, after approximately 10 years of tenure, the promotion probability of internal apprenticeship graduates becomes significantly higher than the promotion probability of workers who did not obtain an apprenticeship degree. Workers with no apprenticeship degrees face significantly lower promotion probability than workers with internal apprenticeship degrees, but there seems to be no significant difference to individuals without an apprenticeship degree and those with external degrees. One reason for our outcomes might be that workers with an internal apprenticeship degree might be more able than workers with an external apprenticeship degree and more able than workers with no degree at all. As the firm is able to gain information about the workers while they undertake internal apprenticeship training, the firm might be able to select workers based on very precise information. In other words, during the three or four years of apprenticeship training, the firm can monitor its workers very precisely. After the apprentices graduate, the firm can easily dismiss low-ability workers because workers and firms in Germany have to negotiate a new working contract after the end of the apprenticeship training. Another reason might be that firm-specific human capital plays a major role within that particular firm, and internal apprentices have acquired a greater amount of this capital.



Note: Own calculations. The solid line represents the predicted values from a locally weighted regression with bandwith 0.8.

Figure 1.4. **a** Promotion and tenure—different profiles from locally weighted regressions (no apprenticeship vs. external apprenticeship). **b** Promotion and tenure—different profiles from locally weighted regressions (no apprenticeship vs. internal apprenticeship)

However, it would be of major interest to know whether different types of workers do indeed differ in their productivity and how their productivity evolves over the course of their careers. Therefore, we complement the analysis with further information on workers' productivity. If we assume productivity-related promotions and wages, our results may imply that the productivity and motivations of workers are negatively associated with age and tenure. Moreover, ability might affect careers through its effect on productivity. If workers of higher ability are also more productive, ability determines careers through workers' performance. To test this implication, we observe tenure- and age-related productivity. As a proxy for productivity, we use information on workers' contributions of suggestions to the company. Suggestions represent workers' efforts to enhance the production process of the company and to increase the overall productivity of the workforce. Figure 1.5 presents workers' likelihoods of making rewarded suggestions with respect to age and tenure, using robust weighted regressions at each observation point.



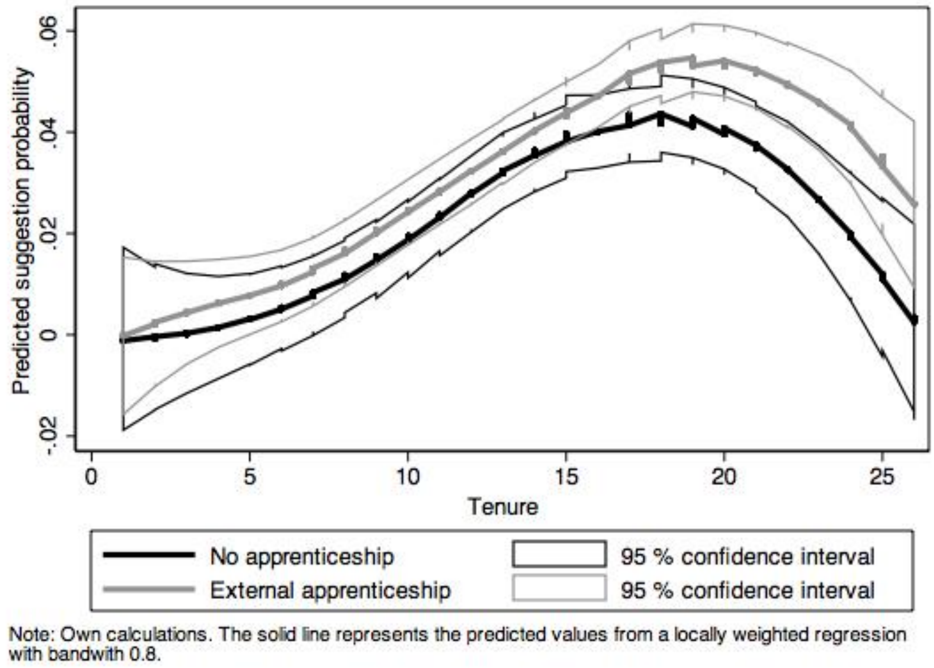
Note: Own calculations. The solid line represents the predicted values from a locally weighted regression with bandwith 0.8.

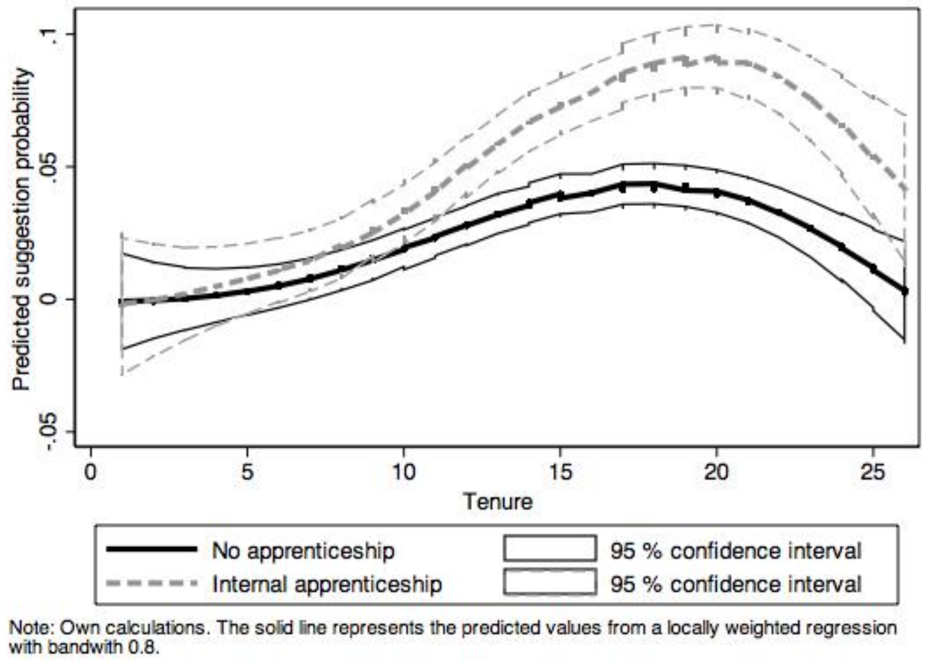
Figure 1.5. Suggestion tenure/age from locally weighted regressions

The probability of making a suggestion exhibits an inverted u-shape with respect to both years of tenure and age. Therefore, a worker is most likely to make a suggestion at middle age, between 36 and 41 years old. Workers who have spent 16-20 years in the company provide the majority of suggestions. Increasing productivity during the beginning of careers may reflect the acquisition of firm-specific capital and learning between the worker and the firm in that the firm is more able to observe the actual productivity of the workers. Declining productivity during the end of careers might be rooted in health- or technology-related human capital depreciation. Generally, this result provides a rationale for why promotions are less likely for elderly workers. If the promotion decision depends on productivity measures, older workers are less likely to be promoted due to decreasing productivity at the end of their careers.

We provide further insight into the connection between productivity and tenure by distinguishing workers with respect to their ability using apprenticeship profiles to proxy ability. Figure 1.6a shows the relationship between suggestions and tenure for external apprentices and for workers who did not obtain an apprenticeship degree.

The inverted u-shaped relation between tenure and suggestion probability remains for all workers, regardless of apprenticeship. However, we cannot find a significant difference between external apprentices and workers who did not obtain an apprenticeship. In Fig. 1.6b, we compare internal apprentices with workers without apprenticeship degrees.





Note: Own calculations. The solid line represents the predicted values from a locally weighted regression with bandwith 0.8.

Figure 1.6. **a** Suggestion tenure—different profiles from locally weighted regressions (no apprenticeship vs. external apprenticeship). **b** Suggestion tenure—different profiles from locally weighted regressions (no apprenticeship vs. internal apprenticeship)

With regard to productivity, we find that workers with an internal apprenticeship degree make significantly more suggestions between 10 and 25 years of tenure. This result might confirm that promotions are productivity related and that firm-specific human capital plays a substantial role in the firm.

Because suggestions are only a proxy for productivity, and the formal apprenticeship is only a proxy for unobserved ability, there might be other explanations for the relation between suggestions and formal apprenticeship qualifications. As technological and/or organizational innovations make companies more efficient, low-qualified workers, in particular, might have an interest in avoiding such innovations to save their own jobs. Consequently, they might avoid improving the production process by making suggestions.

Suggestion probability is slightly higher for externals than for workers without degrees until 16 years of tenure, at which point the difference increases. Internals are much more likely to make suggestions than externals after 10 years of tenure. This result provides further support for productivity-related promotion and wages. Internals face better promotion chances because they show higher productivity compared to workers with external apprenticeship degrees or no degrees.

1.5 Conclusion

This paper aims at understanding the career and productivity developments associated with years of age and tenure in internal labor markets. Using long and balanced panel data on one firm, we find that promotion probability is initially high and decreases significantly with years of both age and tenure. We are able to provide an explanation for lower promotion probability at the end of career based on the lower productivity of older workers. Including ability in the analysis, which we proxy through workers' apprenticeship degrees, offers further implications. First, workers with degrees have a higher promotion probability than workers without, and workers with internal degrees are more likely to make rewarded suggestions to enhance the production process. Second, workers who have received an internal apprenticeship obtain higher probabilities for promotions in hierarchies at the end of careers than workers with external degrees. Furthermore, workers with internal degrees are much more likely to make suggestions than workers with external degrees.

The major contribution of our study is that we are able to describe the career and productivity patterns of workers based on very accurate data over a long time period. This approach has the advantage that our results do not suffer from common empirical problems, such as measurement error or sample attrition. However, these advantages come at the cost of relying on a very selective sample of long-tenured blue-collar workers. Specifically, because of data restriction, we are not able to observe the entire workforce of the firm (including both blue- and white-collar workers), nor do we have information about workers who left the firm during the observation period. Because long-tenured workers are likely to be more productive than workers who are laid-off after short periods and less productive than workers moving voluntarily to other firms, we cannot claim that our estimates represent the average effects for the entire workforce of the firm. Nevertheless, the turnover in German manufacturing firms during the 1980s and 1990s was typically very low and was not comparable with turnover in other countries, such as, for example, the US. Consequently, we do not expect that our results would be substantially different if we were able to observe the entire workforce.

However, given the descriptive nature of our study, our findings still provide first insights with potential important policy implications for the employability of elderly workers. With the ongoing demographic developments, one important question involves measures that might reduce employment barriers for older workers. For our special case, we observe that productivity and promotions both decrease with age. The actual determinant for decreasing productivity, either lower motivation due to lower promotion probability or human capital depreciation, is ambiguous. However, in both cases, human capital investments at later stages of careers might serve to enhance the employability of older workers by increasing the productivity of the elderly workforce. A higher level of firm-specific human capital and a larger productivity level could induce employers to increase wages for the elderly workforce and to extend working contracts, which would then enhance the motivation of older workers.

A detailed description of the adequate instruments needed to alleviate employment barriers for the elderly workforce will require more research. We have shown the importance of analysing internal labor markets. Balanced panel data from personnel records of sufficient length are vital for the identification of the causal effects of age and tenure on careers. We hope for more studies to be performed with similarly long panels and sound information from other companies.

References for Chapter 1

Abowd, J. M. and Kramarz, F. (1999), The analysis of labor markets using matched employer-employee data, *Handbook of Labor Economics* 3(2), 2630–2672.

Akerlof, G. A, Yellen, J. L. (1986), Efficiency wage models of the labor market. Cambridge University Press, Cambridge

Baker, G., Gibbs, M, Holmstrom, B. (1994a), The internal economics of the firm: evidence from personnel data, *Quarterly Journal of Economics* 109(4), 881–920.

Baker, G., Gibbs, M, Holmstrom, B. (1994b), The wage policy of a firm, *Quarterly Journal of Economics* 109(4), 921–956.

Becker, G. S. (1962), Investments in human capital: a theoretical analysis, *Journal of Political Economy* 7(5), 9–49.

Becker, G. S., Stigler GJ (1974), Law enforcement, malfesance and compensation of enforcers, *Journal of Legal Studies* 3, 1–18.

Börsch-Supan A. and Weiss, M. (2007), Productivity and the age composition of work teams: evidence from the assembly line, Mea working paper no. 148.

Chan, W. (1996) External recruitment versus internal promotion, *Journal of Labour Economics* 14(4), 555–570.

Cleveland, W.S. (1979), Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74(368), 829–836.

Cleveland, W. S. (1981), LOWESS: a program for smoothing scatterplots by robust locally weighted regression, *American Statistical Association* 35(1), 54.

Cox, N. J. (2005), Speaking stata: smoothing in various directions, Stata Journal 5, 574–593.

Daveri F. and Maliranta, M. (2006), Age, seniority and labour costs: lessons from the finnish IT revolution, *Economic Policy* 22(49), 117–175.

Dostie, B. (2006), Wage productivity and ageing, *De Economist* 159(2), 139-158.

Fair, R. (1994), How fast do old men slow down, *Review of Economic and Statistics* 76(1), 103–118.

Fair, R. (2007), Estimated age effects in athletic events and chess, *Experimental Aging Research* 33, 37–57.

Fairburn, J.A. and Malcomson, J. M. (2003), Performance, promotion, and the peter principle, *Review of Economics Studies* 68(1), 44–66.

Ford, C. M. (1996), A theory of individual creative action in multiple domains, *Academy of Management Review* 21(4), 1112–1142.

Frank, E. and Nüesch, S. (2011), The effect of wage dispersion on team outcomes and the way team outcome is produced, *Applied Economics* 43, 3037–3049.

Freeman, S. (1977), Wage trends as performance displays productive potential: a model and application to academic early retirement, Bell Journal of Economics 8(2), 419–433.

Frese, M., Teng, E., Wijnen, C. J. (1999), Helping to improve suggestion systems: predictors of making suggestions in companies, *Journal of Organizational Behavior* 20, 1139–1155.

Gibbons, R. and Waldman, M. (2006), Enriching a Theory of Wage and Promotion Dynamics inside Firms, *Journal of Labor Economics* 24(1), 59-107.

Gibbons, R. and Waldman, M. (1999a), Careers in Organizations: Theory and Evidence, in Ashenfelter, O. and Card, D. E. (eds.), Handbook of Labor Economics, vol. 3B, Amsterdam: North-Holland, 2373-2437.

Gibbons, R. and Waldman, M. (1999b), A Theory of Wage and Promotion Dynamics Inside Firms, *Quarterly Journal of Economics* 114(4), 1321-1358.

Goebel, C. and Zwick, T. (2009), Age and Productivity- Evidence from Linked Employer Employee Data. ZEW Discussion Papers 09-020.

Grip, A. D. and Loo, J. V. (2002), The economics of skills obsolescence: a review, *Research in Labor Economics* 21, 1-26.

Grund, C. (2002), Empirische Analyse auf Basis betrieblicher Personaldaten, *WIST* 11, 627-632.

Habermalz, S. (2006), The Speed of Employer Learning and Job Market Signaling Revisited, IZA Discussion Paper No. 2309.

Hellerstein, J. K. and Neumark, D. (2004), Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set NBER Working Paper Series 10325.

Hellerstein, J. K., David, N. and Troske, K. R. (1999), Wages, productivity and worker characteristics: Evidence from plant level production function and wage equations. *Journal of Labor Economics* 17(3), 409-446.

Herpen, M., Cools, K., and Van Praag, M. (2006), Wage Structure and the Incentive Effects of Promotion, *KYKLOS* 59(3), 441-459.

Jacoby, W. G. (2000), Loess: a nonparametric, graphical tool for depicting relationships between variables, *Electoral Studies* 19, 577-613.

Jones, B. F. (2010), Age and great invention, Review of Economic and Statistics 92(1), 1-14.

Kahn, C. and Huberman, G. (1988), Two-sided Uncertainty and “Up-or-Out” Contracts, *Journal of Labor Economics* 6(4), 423-444.

Lazear, E. P. (2000), Performance Pay and Productivity, *American Economic Review* 90(5), 1346-1361.

Lazear, E. P. (1992), The Job as a Concept. In: Bruns, W.J. (ed.), Performance Measurement, Evaluation, and Incentives, Boston, 183-215.

Lazear, E. P. (1981), Agency, Earning Profiles, Productivity and Hours Restrictions, *American Economic Review* 71(4), 606-620.

Lazear, E. P. (1979), Why is there mandatory retirement?, *Journal of Political Economy* 87(6), 1261-1284.

Lazear, E. P. and Oyer, P. (2007), Personnel Economics. Forthcoming in Gibbons, R. and Roberts, J. (eds.), Handbook of Organizational Economics.

Lin, M. J. (2005), Opening the Black Box: The Internal Labor Markets of Company X, *Industrial Relations* 44(4), 659-706.

McCue, K. (1996), Promotion and Wage Growth, *Journal of Labor Economics* 14(2), 175-209.

McLaughlin, K. J. (1994), Rigid Wages? *Journal of Monetary Economics* 34(3), 383-414.

Medoff J. and Abraham K. (1981), Are Those Paid More Really More Productive?, *Journal of Human Resources* 16(2), 186-216.

Medoff J. and Abraham K. (1980), Experience, Performance, and Earnings, *Quarterly Journal of Economics* 95(4), 703-736.

Mincer, J. (1974), Schooling, Experience, and Earnings. New York. NBER Press.

Murphy, K. J. (1986), Incentives, Learning and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts, Rand Journal of Economics 17(1), 59-76.

OECD (2006), Live Longer, Work Longer. OECD Paris.

Oster, S. M. and Hamermesh, D. S. (1998), Ageing and Productivity Among Economists, *The Review of Economics and Statistics* 80(1), 154-156.

Prendergast, C. (1993), The Role of Promotion in Inducing Specific Human Capital Acquisition, *Quarterly Journal of Economics* 108(2), 523-534.

Royston, P. (1991), Lowess smoothing. Stata Technical Bulletin 3: 7–9, Reprinted in Stata Technical Bulletin Reprints, vol. 1, pp. 41–44. College Station, TX: Stata Press.

Royston, P., and Cox, N. J. (2005), A multivariable scatterplot smoother, *Stata Journal* 5, 405–412.

Schöttner, A. and Thiele, V. (2010), Promotion Tournaments and Individual Performance Pay, *Journal of Economics and Management Strategy* 19(3), 699-731.

Skirbekk, V. (2003), Age and Individual Productivity: A Literature Survey, MPIDR Working Paper WP 2003-028.

Treble, J., van Gameren, E., Bridges, S., and Barmby, T. (2001), The Internal Economics of the Firm: Further Evidence from Personnel Data, *Labour Economics* 8(5), 531-552.

Van Ours, J. and Stoeldraijer, L. (2010), Age, Wage and Productivity, IZA Discussion Papers 4765.

Waldman, M. (2007), Theory and evidence in internal labor markets, MPRA Discussion Paper No. 51.

Weinberg, B. A. and Galenson, D. W. (2005), Creative careers: the life cycles of Nobel laureates in economics, NBER working paper no. 1179.

Appendix for Chapter 1

See Table 1.2

Table 1.2. Results from regression analysis

	Wage group		Promotions		Suggestion	
	(1) OLS	(2) OLS	(3) LPM	(4) LPM	(5) LPM	(6) LPM
Age (Ref: age 1519)						
Age 2024	1.597*** (0.104)		−0.039** (0.020)		0.008** (0.003)	
Age 2529	2.365*** (0.109)		0.108*** (0.017)		0.024*** (0.006)	
Age 3034	3.200*** (0.123)		0.126*** (0.017)		0.053*** (0.009)	
Age 3539	3.814*** (0.137)		0.152*** (0.017)		0.075*** (0.009)	
Age 4044	4.247*** (0.152)		0.151*** (0.016)		0.060*** (0.009)	
Age 4554	4.481*** (0.171)		0.172*** (0.019)		0.034*** (0.009)	
Tenure (Ref: tenure 0105)						
Tenure 0610		0.583*** (0.045)		0.060*** (0.008)		0.001 (0.004)
Tenure 1115		1.321*** (0.075)		0.066*** (0.008)		0.032*** (0.007)
Tenure 1620		1.951*** (0.096)		0.089*** (0.007)		0.059*** (0.007)
Tenure 2127		2.421*** (0.116)		0.082*** (0.007)		0.027*** (0.006)
Schooling (Ref: no/low degree)						
Higher school degree	2.533*** (0.640)	2.749*** (0.608)	0.042*** (0.011)	0.037*** (0.011)	0.028*** (0.008)	0.025*** (0.008)
Apprenticeship (Ref: no degree)						
Internal apprenticeship	2.840*** (0.218)	3.085*** (0.223)	0.038*** (0.006)	0.029*** (0.006)	0.028*** (0.011)	0.031*** (0.011)
External apprenticeship	0.767*** (0.155)	1.388*** (0.156)	0.036*** (0.005)	0.011** (0.005)	0.003 (0.007)	0.010 (0.006)
Constant	2.475*** (0.096)	3.938*** (0.091)	0.165*** (0.015)	0.121*** (0.005)	0.013*** (0.004)	0.003 (0.005)
Random effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.385	0.341	0.031	0.019	0.021	0.020
χ^2 value	1145.59	755.88	361.80	270.66	98.27	76.24
Number of observations	10544	10544	10129	10129	10544	10544
Number of workers	415	415	415	415	415	415

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Dependent variable are wage group (mean = 6.730), promotion (mean = 0.083), and suggestion (mean = 0.035)

Chapter 2

Training participation of a firm's aging workforce*

Joint work with

Christian Pfeifer, Simon Janssen, Uschi Backes-Gellner

*This chapter is based on the publication Pfeifer, C., Janssen, S., Yang, P., and Backes-Gellner, U. (2013), Training participation of a firm's aging workforce, *Empirical Research in Vocational Education and Training* 4(2), 131-147.

2.1 Introduction

Older workers have on average higher employment stability than younger workers but lower reemployment probabilities and often longer unemployment durations in most countries (e.g. Hutchens, 1988; Chan and Stevens, 2001; OECD, 2005; OECD, 2008; EU, 2009). This leads to hardships for unemployed older workers (e.g. loss in consumption standards, psychological burden due to loss of main activity and social networks) and to society, because tax payers have to finance unemployment benefits or early retirement schemes. Therefore, identifying the factors that might lead to employment barriers for older workers is of central importance in times of demographic change. One major economic explanation for employment barriers is that older workers might have a productivity that is lower than their wages. As productivity is largely determined by human capital investments, the relationship between training and aging can help us to understand disadvantages of older workers in the labor market. If firms and workers invest less in human capital at later stages of workers' careers, and if firms cannot adjust individual wages (e.g. due to collective contracts or minimum wage legislations), it would be less profitable for firms to employ older workers. Productivity enhancing training might alter the incentives to expand employment contracts in current firms and to integrate older unemployed workers in new firms (Gruber and Wise, 1998). In times of rapid technological change, training becomes increasingly important because computer-based technologies demand a new range of abilities, which older workers need to acquire in order to avoid the depreciation of skills and competencies (Friedberg, 2003).

In this paper, we analyze the impact of aging on the participation probability in employer-financed training in an internal labor market to shed some light into the black box of training decisions in firms. For this purpose, we develop a simple model for the timing decision when to train a worker, which accounts for screening and amortization period effects and from which our econometric framework is generated. We further use a personnel data set that contains information on more than 10,000 yearly observations for 400 male blue-collar workers of a German company for four entry cohorts. The length of the panel is longer than 24 years. The data contain unique information about four different training measures: short training course, longer training course, longer vocational re-training, and academy of vocational training. Although results from personnel data are not necessarily representative, they have the advantage of overcoming unobserved firm and training course heterogeneity, which might bias results from survey data. To analyze the effect of age on training participation, we apply random effects Logit and multinomial Logit regressions. The main results of our econometric case study are that training participation is inverted u-shaped with age and that longer training is performed earlier in life. These findings are in line with predictions from our theoretical model.

The subsequent paper is structured as follows. The next section summarizes previous findings on age and training participation. In Section 3, we present a model for timing of training participation, from which we generate our main research hypotheses and estimation framework. Section 4 informs about the personnel data set and provides descriptive statistics. The regression analyses are presented in Section 5. The paper concludes with a short summary and a discussion of the results.

2.2. Previous research

Several empirical studies analyze worker characteristics to explain individual variation in training participation and find that education and age are the most important factors. Frazis et al. (2000) draw from a rich database of employer and employee surveys to analyze the educational effect on training participation in the U.S. They find significant positive effects of educational attainment on the incidence and intensity of formal training. Similar results are found in panel data of young U.S. workers (Veum, 1997) and in European data (Oosterbeek, 1998; Arulampalam et al., 2004; Arulampalam and Booth, 1997).

Theoretical models with respect to age and training emphasize two main arguments: the amortization (payback) period of training investments and the signaling function of training. The former explanation states that older workers are less likely to receive training due to lower total net returns associated with shorter time horizons until retirement (Becker, 1962; Becker, 1993). Therefore, the investments into older workers have to yield significantly larger gains to make their training profitable, especially when facing deferred payment schemes (Lazear, 1979). The signaling function of training refers to information asymmetries. After incurring hiring costs, firms still know little about the potential ability and productivity of the new employees. Training might reduce information asymmetries and is most effective early in workers’ careers (Acemoglu and Pischke, 1998). Overall, both arguments (amortization period and signaling) predict a negative correlation between training and age.

Oosterbeek (1998) uses Dutch household panel data to estimate univariate and bivariate Probit models with linear age as explanatory variable. He finds small but significant negative age effects on training. Maximiano and Oosterbeek (2007) evaluate the impact of age on workers’ willingness to receive training and employers’ willingness to provide training. They also report a small but significant negative linear age effect. Studies with non-linear age specifications provide a more detailed view of the correlation between age and training incidence. Leuven and Oosterbeek (1999) include binary age categories as independent variables in Probit and linear probability models of training. The results are heterogeneous with respect to size, direction, and significance across different countries. Whereas Canada and the Netherlands suggest an inverted u-shaped relation between age and training probability, Switzerland and the U.S. reveal no significant association. O'Connell and Byrne (2009) extend the empirical investigations by controlling for binary age categories within a multinomial Probit regression. Training classification distinguishes between no training, general training, and specific training. The empirical results suggest an inverted u-shaped relationship between age and training, which exhibits weak robustness when including further control variables. An inverted u-shaped age curve for participation in training is also found by Sousa-Poza and Henneberger (2003), who use age, squared and cubed age as explanatory variables for training probability. The results provide small but robust age effects. Riphahn and Trübswetter (2006) also find an inverted u-shaped association between age and training in German microcensus data.

Whereas the downward-sloping part of the inverted u-shaped relationship, which has been found in several studies, can be explained by amortization period and signaling effects, the upward-sloping part cannot. We therefore develop a new simple model for the timing decision of training participation in the next section, from which we derive our econometric framework and research hypotheses.

2.3 A model for timing of training participation

The focus of our subsequent model about age and training participation is not on the question of whether a firm and a worker invest in human capital, which is the core of most models, but on the question of when the investment is undertaken. For simplicity, we do not distinguish between firm and worker decisions; instead we treat training as a joint decision.[13] As we discuss the effects on total rents, the rent-sharing aspect of human capital investments can be

[13] In principal, workers and firms face the same effects. Thus, we would obtain the same insights if the investment decisions were analyzed separately. An advantage of analyzing the joint decision is that we can neglect the rent-sharing aspect of human capital investments.

neglected and, hence, wages do not need to be incorporated into our model. Moreover, human capital investment is a binary choice variable, because our paper is about participation in training courses.

The basic mechanisms in our model are a «screening/learning effect» and an «amortization period effect» which have different directions. Younger workers are more engaged in job shopping and firms have to undertake more screening of younger workers, because uncertainty of their quality and willingness to stay in a specific job and firm is higher (e.g. Jovanovic, 1979; Topel and Ward, 1992; Farber and Gibbons, 1996; Lange, 2007). Consequently, firms and workers have less incentives to invest in (firm-specific) human capital at the start of an employment relationship when a worker is young. If the match between worker and firm proves to be of good quality, both parties have incentives to undertake human capital investments. The worker benefits from higher future wages due to higher future productivity and from signalling to the firm higher productivity and work attachment, which increases his promotion probability and long-term income prospects. The firm benefits from higher future productivity of workers. The firm furthermore might need some time to learn about workers' skills to determine training contents and to select participants. We therefore expect that the training participation probability is positively correlated with age for younger workers («screening/learning effect»). Investment incentives, however, decline with age because the amortization period decreases as a worker gets older and approaches retirement («amortization period effect»). While the total effect of age on training should be dominated by the «screening/learning effect» in the first years of workers' careers, the «amortization period effect» should dominate thereafter.

Let us now turn to the simple model. The decision to train a worker depends on total net rents of training *R* in equation (1).[14] The net rents are the total increase in the value of productivity due to training ΔP (compared to the situation in which a worker receives no training) over all years *t* until retirement is reached minus the total fixed costs *C* of the training course. The age at which training takes place is denoted by *a* and retirement age by r. The length of the amortization period in years is therefore *r–a*. In the subsequent discussion, we consider two cases. The first case assumes no depreciation of human capital acquired in the training course, which leads to a constant productivity increase over time ($\Delta P_t = \Delta P_0$), while the second case acknowledges human capital depreciation.

$$R[a] = \Delta P - C = \sum_{t=1}^{r-a} \Delta P_t - C \tag{1}$$

We begin by illustrating the «amortization period effect» for the first case. The total net rent is depicted in equation (2) and its first derivate with respect to age in equation (3). We see that one more year of age at training, which implies a reduction of the amortization period by one year, decreases the total net rent linear by the foregone higher value of productivity in that additional year.

$$R = \Delta P - C = \sum_{t=1}^{r-a} \Delta P_0 - C = (r-a)\Delta P_0 - C \tag{2}$$

$$\frac{\partial R}{\partial a} = -\Delta P_0 \tag{3}$$

In the next step, we introduce the «screening/learning effect». The «screening/learning effect» implies that the productivity increase is to some extent uncertain, which is represented through the expected total productivity increase as presented in equation (4). Firms and

[14] Table A.1 in the Appendix contains a list of the model's variables.

workers need to learn about the match quality and the willingness to engage in a long-term contract to benefit from returns of human capital investments. The firm further needs to learn about a worker's human capital stock to determine course contents. Both learning necessities can be introduced through a learning parameter γ, which is a non-linear function of worker's age at training. If training takes place later in life, the more has been learned about a worker, but with decreasing marginal returns to learning.[15] Because the learning parameter γ is restricted to values between zero and one, γ can be interpreted as the probability that a worker has the increased productivity after training and $(1-\gamma)$ as the probability that training does not increase productivity ($\Delta P = 0$).

$$E[\Delta P] = \gamma[a]((r-a)\Delta P_0) with \frac{\partial \gamma}{\partial a} > 0, \frac{\partial^2 \gamma}{\partial a^2} < 0 \tag{4}$$

Equation (5) presents the expected total net rent combining the «amortization period effect» and the «screening/learning effect». The first derivate in equation (6) shows that the expected total net rent increases with age as long as $\partial\gamma/\partial a\ (r-a)\ \Delta P_0 > \gamma\,[a]\ \Delta P_0$ and decreases with age if $\partial\gamma/\partial a\ (r-a)\ \Delta P_0 < \gamma\,[a]\ \Delta P_0$. It can be seen that the left hand side of the first order condition for the maximum expected total net rent in equation (7) decreases with age and that the right hand side increases with age.

This is also reflected in the second derivate in equation (8). Overall, the age effect is non-linear with an inverted u-shaped relationship between the expected total net rent of training and the age at which training takes place.

$$E[R] = E[\Delta P] - C = \gamma[a]((r-a)\ \Delta P_0) - C \tag{5}$$

[15] Note that learning in our model depends only on age. This can be reasoned by the fact that workers in our model are homogeneous with respect to entry age and tenure is age minus entry age. A rationale in a model with heterogeneous entry age would be that learning can also take place through previous work careers in other firms (e.g. experience, signals).

$$\frac{\partial E[R]}{\partial a} = \underbrace{\frac{\partial \gamma}{\partial a}(r-a)\Delta P_0}_{>0} - \underbrace{\gamma[a]\Delta P_0}_{>0} = 0 \tag{6}$$

$$\underbrace{\frac{\partial \gamma}{\partial a}(r-a)\Delta P_0}_{a\uparrow\Rightarrow\downarrow} = \underbrace{\gamma[a]\Delta P_0}_{a\uparrow\Rightarrow\uparrow} \tag{7}$$

$$\frac{\partial^2 E[R]}{\partial a^2} = \underbrace{\frac{\partial^2 \gamma}{\partial a^2}(r-a)\Delta P_0}_{<0} - 2\underbrace{\frac{\partial \gamma}{\partial a}\Delta P_0}_{>0} < 0 \tag{8}$$

We now consider the second case with human capital depreciation, which leads qualitatively to same results as the first case. Human capital depreciation is introduced through the depreciation factor $(1 + \beta)^t > 1$, i.e. the productivity increase due to training is lower in later periods than in earlier periods after training participation ($\Delta P_t = \Delta P_0(1 + \beta)^{-t}$). The new expected total net rents from training are presented in equation (9). From the first derivate in equation (10) and the second derivate in equation (11), we can again see that the relationship between expected total net rents and age at training is also inverted u-shaped if we account for human capital depreciation.

$$E[R] = \gamma[a]\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t} - C \tag{9}$$

$$\frac{\partial E[R]}{\partial a} = \underbrace{\frac{\partial \gamma}{\partial a}\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t}}_{>0} - \underbrace{\gamma[a]\frac{\Delta P_0 \ln(1+\beta)}{(1+\beta)^{(r-a)}\beta}}_{>0} \stackrel{!}{=} 0 \tag{10}$$

$$\frac{\partial^2 E[R]}{\partial a} = \underbrace{\frac{\partial^2 \gamma}{\partial a^2}\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t}}_{<0} - 2\underbrace{\frac{\partial \gamma}{\partial a}\frac{\Delta P_0 \ln(1+\beta)}{(1+\beta)^{(r-a)}\beta}}_{>0} - \underbrace{\gamma[a]\frac{\Delta p_0 \ln(1+\beta)^2}{(1+\beta)^{(r-a)}\beta}}_{>0} < 0 \tag{11}$$

The probability of participating in training at a given age (T_a=1) is depicted in equation (12) and depends on expected total net rents at that age. To be more precise, training takes place (T_a=1) if the expected total net rents plus an idiosyncratic normally distributed error term ε with zero mean are larger than some threshold value z. Because we have shown that expected total net rents are inverted u-shaped with age, the training probability should also be inverted u-shaped with age.

$$\Pr [T_a = 1| E [R [a]]] = \Pr [E [R [a]] + \varepsilon > z] \tag{12}$$

From equation (12), we can derive our econometric model applying a second order Taylor approximation to the expected total net rents *(E* [*R*]). Equation (13) states the basic Logit model we have to estimate, in which P_1 and p_2 denote the coefficients for age and squared age, λ are the coefficients for a vector of control variables X, and $\wedge$ is the cumulative density function of the logistic distribution.

$$\Pr [T_a = 1|a,X] = \Pr [p_1 a + p_2 a^2 + \lambda X + \varepsilon > z] = \wedge[p_1 a + p_2 a^2 + \lambda X] \tag{13}$$

To summarize, we can formulate our main research hypothesis on the timing of training, which is then tested using longitudinal personnel data and Logit models in the next sections.

Hypothesis 1: The training participation probability is inverted u-shaped with age ($p_1 > 0$ and $p_2 < 0$).

Our model also allows us to generate an additional hypothesis. Longer and, consequently more expensive, training courses are likely to increase productivity *(ΔP)* by more than shorter training courses. Therefore, the «amortization period effect» *($-\Delta P_t$)* is larger for longer training courses so that expected net rents are, ceteris paribus, maximized at earlier training age.

Hypothesis 2: The training participation probability peaks at earlier age for longer training courses.

2.4 Data set and descriptive statistics

We use personnel data of a large German company from the energy sector located in West Germany. The company is subject to a collective contract and has a works council. Due to data protection reasons, we are neither allowed to name the company nor to provide any further information. The data contain a subsample of 438 blue-collar workers in the company's mining business. All of these workers entered the firm in four subsequent cohorts, from 1976 until 1979, and stayed in the company over the entire observation period up to the year 2002. The sample represents a share of about 25 percent of all employees in the company's operation unit and 3.5 percent of the company's entire workforce.

A disadvantage of our quasi-balanced panel design is that we have no information about workers who left the firm so that we cannot control for a potential selection bias. The data set is nevertheless adequate to study the long-term issues of an aging workforce and of career aspects in the context of human capital investments due to its large panel length. We include only German male blue-collar workers without missing values in the variables we use. This restriction reduces our sample by only 5 percent. The final sample contains 10,544 yearly observations of 415 different workers (1976: number of workers n = 105, panel length in years T = 27; 1977: n = 96, T = 26; 1978: n = 77, T = 25; 1979: n = 137, T = 24).

The data set allows us to use two kinds of training variables. The first variable is binary and takes the value one if a worker participated in training in a given year. Thus, we can apply a random effects Logit model. The second variable indicates what kind of training a worker received so that multinomial Logit models are appropriate. If a worker did not participate in training in a given year the value is zero. For training participation, we have information about four different training measures: (1) short training course («kurze Schulung») (one or two days), (2) longer training course («längere Schulung») (up to several weeks), (3) longer vocational re-training («längere Umschulung») (up to several weeks), and (4) longer academy of vocational training («Berufsakademie») (up to several weeks). Unfortunately, we do not have information about earnings of workers. We know however that workers are paid during the training measures and do not have to cover any direct costs. Table 2.1 presents summary statistics of the training measures. On average 6.3 percent of the workers in our sample participated in some kind of training in an average year, which results in 664 training cases in our observation period. About two thirds of all cases are short training courses, whereas the other training measures are nearly equally distributed.

Table 2.1. Descriptive statistics of training variables

	Mean	**Standard deviation**	**Training cases (total number)**
Training (all) (dummy)	0.0630	0.2429	664
Training measures (reference (0) no training):			
(1) Short training	0.0405	0.1971	427
(2) Longer training	0.0073	0.0851	77
(3) Longer re-training	0.0068	0.0824	72
(4) Longer academy	0.0083	0.0910	88

Notes: Number of observations is 10,544 from 415 workers in a balanced panel design.

Our main interest lies in the impact of age on training participation. We specify age in two non-linear ways in the subsequent regression analyses. First, we use dummy variables for the age category. Second, we use age in years and its higher terms. Though most age variance stems from within as we observe workers for at least 24 years, between-age variance also exists as the workers were born between 1952 and 1963. We further consider dummy variables for schooling and apprenticeship degrees to account for skill differences of workers at the time they enter the firm. More information about the explanatory variables is given in Table A.2 in the Appendix.

First descriptive evidence for the impact of age on the overall training participation probability is depicted in Figure 2.1. The results are based on estimations using robust locally weighted regressions. This is a non parametric approach to smooth scatter plots based on multiple weighted linear regressions for every observation point (Cleveland, 1979). It can be seen that our expected inverted u-shape relationship is indeed confirmed by the data, which stresses the importance of non-linear specification of age when estimating the determinants of training participation.

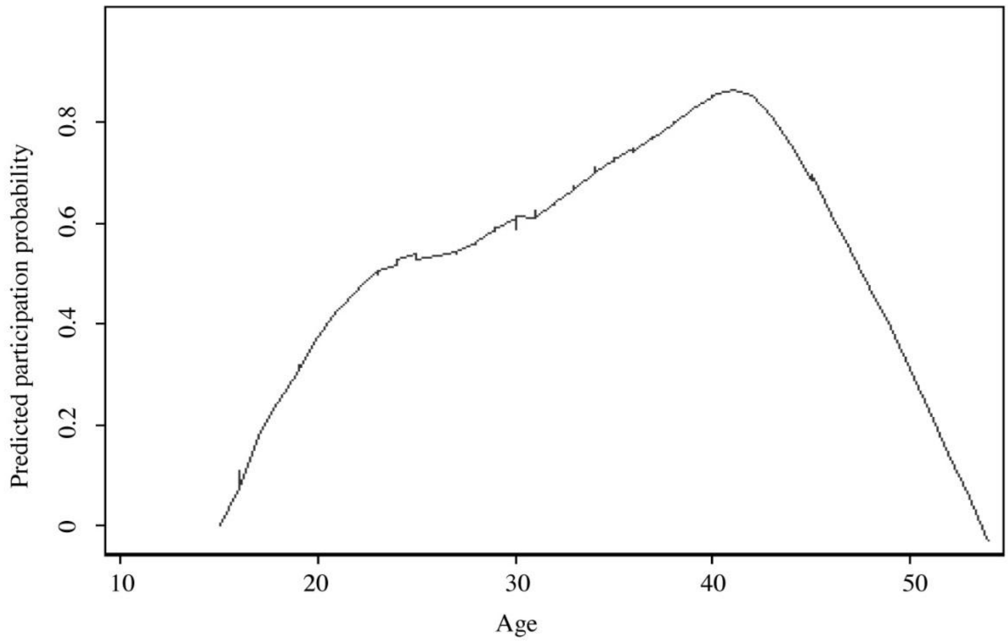


Figure 2.1. Age and participation probability in training from locally weighted regressions

2.5 Regression analyses

At first, we estimate a random effects Logit model for the general participation probability in training. The likelihood ratio test rejects the null hypothesis that the variance of the random effects is zero. As our dependent variable is binary and has a rather low expected probability, linear regressions would yield a high share of outside predictions. We estimate two specifications, which reveal in principal the same results. The first specification includes dummy variables for age categories and the second specification includes polynomials of age in years (until the quartic term). The results of the binary random effects Logit regressions are presented in Table 2.2. Though we also present the coefficients, our main interest is on marginal effects at the means of all covariates as well as on predicted probabilities.

Table 2.2. Determinants of training participation

	(1-Coeff.)		**(2-Coeff.)**		**(1-Mfx)**		**(2-Mfx)**	
Age categories (reference:15–19)								
Age category 20–24 (dummy)	0.974 (0.343)	**			0.061 (0.028)	**		
Age category 25–29 (dummy)	1.149 (0.338)	***			0.073 (0.029)	***		
Age category 30–34 (dummy)	0.854 (0.342)	**			0.050 (0.025)	**		
Age category 35–39 (dummy)	1.699 (0.333)	***			0.129 (0.038)	***		
Age category 40–44 (dummy)	1.643 (0.337)	***			0.129 (0.041)	***		
Age category 45–54 (dummy)	0.759 (0.408)	*			0.047 (0.033)	*		
Age polynomials								
Age in years			9.153 (1.735)	***				
Age squared / 100			–44.873	***				

	(1-Coeff.)		(2-Coeff.)		(1-Mfx)		(2-Mfx)	
			(8.33)					
Age cubed / 1000			9.563 (1.737)	***				
Age quartic / 10000			–0.746 (0.133)	***				
Mfx: Age polynomials							0.001 (0.0005)	**
Schooling (reference: no school degree)								
Low school degree («Hauptschule») (dummy)	0.014 (0.153)		0.014 (0.152)		0.001 (0.007)		0.001 (0.007)	
Higher school degree (at least «Realschule») (dummy)	0.488 (0.227)	**	0.482 (0.226)	**	0.027 (0.015)	**	0.024 (0.011)	**
Apprenticeship (reference: no apprenticeship)								
Apprenticeship degree in firm	0.438 (0.166)	**	0.412 (0.165)	**	0.022 (0.009)	**	0.020 (0.008)	**
Apprenticeship degree outside firm	0.105 (0.144)		0.077 (0.144)		0.005 (0.007)		0.004 (0.007)	
Observations	10544		10544					
Wald test	111.83	***	99.2	***				
LR test of rho=0	65.82	***	65.67	***				

Note: Random effects Logit (coefficients and marginal effects). Standard errors in parentheses. Standard errors for marginal effects are calculated by using the delta method.*** $p<0.01$, **$p<0.05$, *$p<0.10$.

The first specification in Table 2.2 indicates that training participation is inverted u-shaped with age and peaks during the middle-age years, between 35 to 45, which is in line with our first hypothesis. We further use the results of our second specification to plot predicted probabilities in Figure 2.2. The participation probability is to some degree inverted u-shaped with age. As we have considered higher age polynomials, we do not smooth the age effect as we did in the robust locally weighted regressions in Figure 2.1 in the previous section. That we do not find a smoother u-shaped pattern is also reasoned by training course heterogeneity in the binary pooled training measure we employ. Therefore, a multinomial Logit model for different training measures is likely to identify age effects more accurately.

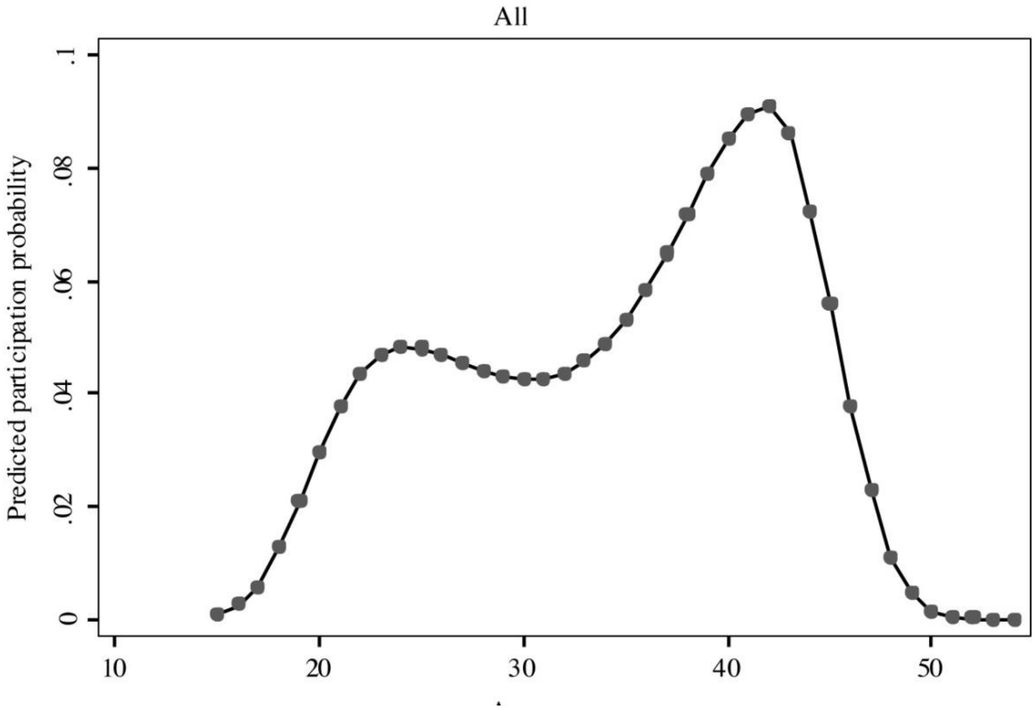


Figure 2.2. Age and predicted participation probability in training from random effects Logit

Our Logit estimates in Table 2.2 further show that workers with higher schooling (at least «Realschule») and workers with an apprenticeship degree earned in the firm have significantly higher training participation probabilities. Differences between low schooling and no school degree as well as between an outside apprenticeship and no apprenticeship degree are not significant. Higher schooling is likely to be associated with higher levels of general human capital, whereas an internal apprenticeship is associated with job specific human capital. Both kinds of human capital might have a self-productivity effect on further skill acquisition, which increases incentives to invest in training for the worker as well as the firm. The firm might also have better knowledge of qualifications and skills of their own former apprentices and can therefore determine training contents and predict outcomes (e.g. training success, productivity effects, and willingness to stay in firm) more precisely.

In the next step, we use a multinomial Logit model to estimate participation probabilities in different training measures [(0) no training, (1) short training course, (2) longer training course, (3) longer vocational re-training, (4) longer academy of vocational training], which includes age polynomials and only dummies for higher schooling (at least «Realschule») and internal apprenticeship absolved within the firm. We leave out the other educational categories, because we would otherwise have the problem of perfect predictions in different outcome variables. As has been shown in the previous binary Logit estimates, the reduction of categories is reasonable because we have not found significant differences between workers without a school degree and workers with the lowest school degree («Hauptschule») and between workers without apprenticeship degrees and workers with apprenticeship degrees earned in other firms.

The multinomial Logit model is often criticized because of its reliance on the independence of irrelevant alternatives (IIA) assumption and some authors argue for using the Probit rather than the Logit approach (e.g. Alvarez and Nagler, 1995). Recent studies show however that the multinomial Logit model performs better in practice, even under serious violations of the IIA (Dow and Endersby, 2004; Kropko, 2008). We decided to use the Logit approach and carried out a test in order to check whether the IIA is violated in our special case. In detail we carried out the test proposed by Hausman and McFadden (1984). The null hypothesis that the odds of our different outcome categories are independent of other alternatives could not be rejected for any category. Table 2.3 informs about the multinomial Logit regression results.

Table 2.3. Determinants of participation in different training measures

	(1) Short training		(2) Longer training		(3) Longer re-training		(4) Longer academy	
Age polynomials								
Age in years	5.414 (1.951)	***	4.188 (6.479)		13.830 (6.901)	**	30.392 (12.897)	**
Age squared / 100	–29.896 (9.427)	***	–17.029 (33.584)		–69.635 (35.672)	*	–131.686 (60.284)	**
Age cubed / 1000	7.077 (1.972)	***	3.000 (7.590)		15.272 (8.037)	*	24.961 (12.362)	**
Age quartic / 10000		***	–0.206 (0.631)		–1.238 (0.666)	*	–1.756 (0.938)	*
Schooling (Reference: no/low degree)								
Higher school degree (at least «Realschule») (dummy)	–0.005 (0.179)		0.662 (0.335)	**	0.108 (0.469)		1.428 (0.239)	***
Apprenticeship (Reference: no/external degree)								
Apprenticeship degree in firm (dummy)	0.227 (0.113)	*	0.430 (0.243)	*	–0.918 (0.360)	**	1.644 (0.234)	***
Constant	–39.833 (14.713)	**	–41.354 (45.986)		–104.887 (49.088)	**	–263.261 (102.144)	***
Observations				10544				
LR Chi²(24)				517.200***				
Pseudo R²				0.082				

Note: Multinomial Logit (coefficients). Standard errors in parentheses. *** p<0.01, **p<0.05, *p<0.10.

To make interpretation of the results in the multinomial Logit model easier, we plotted the predicted probabilities at different age levels for each training measure in Figure 2.3. Short training courses are the most frequently used measure, which peak in probability at age 42. Longer training and re-training courses have quite similar profiles with peaks between 23 and 25 years. Longer training in the academy is most likely to occur in the late 20s. For each training measure, we find an inverted u-shaped impact of age, which is more pronounced than in the previous binary Logit estimates for the pooled training probability. The results further indicate that longer and, hence, more costly training measures are more likely to be undertaken earlier in life, which supports our second hypothesis. Older workers seem to receive only short training to update their skills. Career enhancing training (academy) is mainly performed by middleaged workers and training to close a qualification gap (longer training and re-training) is primarily performed by young workers.

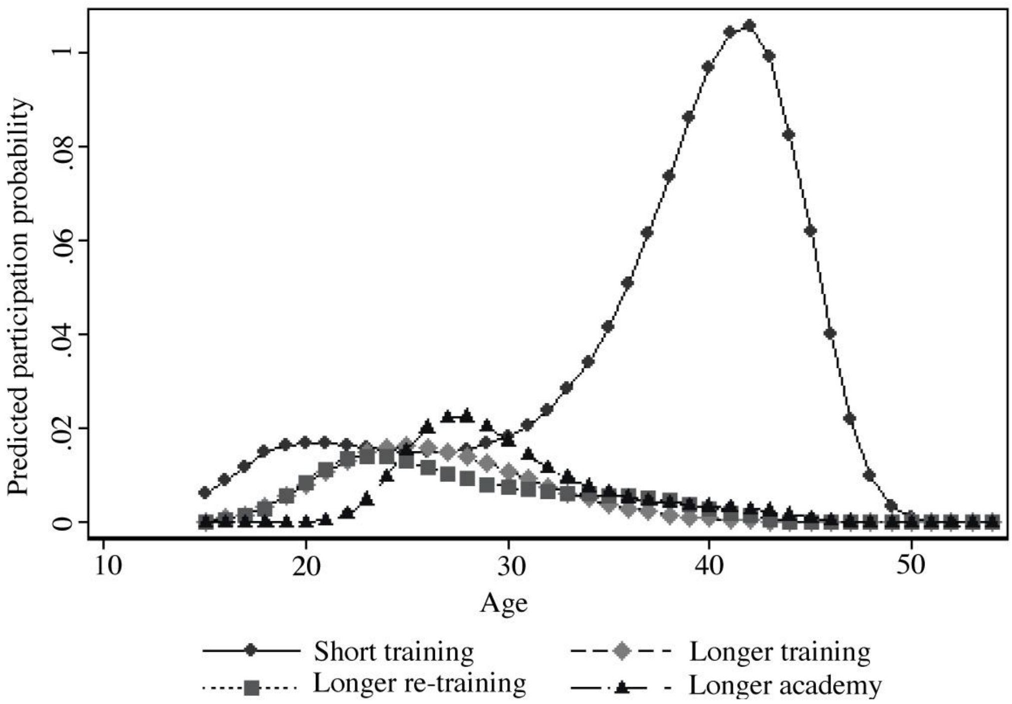


Figure 2.3. Age and participation probabilities of different training measures from multinomial Logit

We can also see from the multinomial Logit results in Table 2.3 that workers with higher secondary schooling are more likely to receive longer training and to attend academy training. Workers with an internal apprenticeship are more likely to receive short and longer training as well as academy training but are less likely to get vocational re-training. The latter result is quite plausible as outside workers might have the wrong qualifications for the job and need re-training. Job-specific and firm-specific skills acquired during an internal apprenticeship might have a self-productivity effect on acquiring further specific skills, which might explain the enormous advantage of insiders in attending the academy for vocational training because in this training measure advanced skills are taught.

2.6 Conclusion

The main results of our econometric case study are that (1) training participation is inverted u-shaped with age, (2) longer training courses are mainly performed earlier in the career, and (3) old workers above the age of 50 years are unlikely to receive any training. Especially the low training probability of older workers, which is likely to be caused by shorter amortization periods, might explain disadvantages of older workers in the labor market (e.g. low re-employment probability). A possible policy intervention are training subsidies targeted at older workers that could counter the effect of decreasing amortization periods and, consequently, should increase the training participation probability, which would hopefully enhance productivity and employability of older workers. Because the amortization period decreases with age, the training subsidies should also increase with age to be effective.

References for Chapter 2

Acemoglu, D. and Pischke, J.-S. (1998), Why do firms train? Theory and evidence, *Quarterly Journal of Economics* 113, 79–119.

Alvarez, M. R. and Nagler, J. (1995), Economics, issues and the Perot candidacy: voter choice in the 1992 presidential election, *American Journal of Political Sciences* 39(3), 714–744.

Arulampalam, W. and Booth, A. L. (1997), Who gets over the training hurdle? A study of the training experiences of young men and women in Britain, *Journal of Population Economics* 10(2), 197–217.

Arulampalam, W.; Bryan, M. L. and Booth, A. L. (2004), Training in Europe, *Journal of the European Economic Association* 2(2–3), 346–360.

Becker, G. S. (1962), Investments in human capital: a theoretical analysis, *Journal of Political Economy* 7(5), 9–49.

Becker, G. S. (1993), Human capital: a theoretical and empirical analysis with special reference to education. 3rd ed. Chicago: The University of Chicago Press.

Booth, A. L. and Bryan, M. L. (2002), Who pays for general training? New evidence for British men and women, IZA Discussion Paper, 486.

Chan, S. and Stevens, A. H. (2001), Job loss and employment patterns of older workers, *Journal of Labor Economics* 19(2), 484–521.

Cleveland, W. S. (1979), Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association,* 74(368), 829–836.

Dow, J. K. and Endersby, J. W (2004), Multinomial probit and multinomial logit: a comparison of models for voting research, *Electoral Studies* 23(1), 107–122.

EU [European Union] (2009), Employment in Europe. Brussels: European Union.

Farber, H. S. and Gibbons, R. (1996), Learning and wage dynamics, *Quarterly Journal of Economics* 111, 1007–1047.

Frazis, H.; Gittleman, M. and Joyce, M. (2000), Correlates of training: an analysis using both employer and employee characteristics, *Industrial and Labor Relations Review* 53(3), 443–462.

Friedberg, L. (2003), The impact of technological change on older workers: evidence from data on computer use, *Industrial and Labor Relations Review* 56(3), 511–529.

Gruber, J. and Wise, D. (1998), Social security and retirement: an international comparison, *American Economic Review* 88(2), 158–163.

Hausman, J. and McFadden, D. (1984), Specification tests for the multinomial logit model, *Econometrica* 52(5), 1219–1240.

Hutchens, R. M. (1988), Do job opportunities decline with age?, *Industrial and Labor Relations Review* 42(1), 89–99.

Jovanovic, B. (1979), Job matching and the theory of turnover, *Journal of Political Economy* 87, 972–990.

Kropko, J. (2008), Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data, Chapel Hill: University of North Carolina [Dissertation].

Lange, F. (2007), The speed of employer learning. *Journal of Labor Economics* 25, 1–35.

Lazear, E. P. (1979), Why is there mandatory retirement?, *Journal of Political Economy* 87(6), 1261–1284.

Leuven, E. and Oosterbeek, H. (1999), The demand and supply of work-related training: evidence from four countries, *Research in Labor Economics* 18, 303–330.

Maximiano, S. and Oosterbeek, H. (2007), On the determinants of workers' and firms' willingness to train, Conference paper, European Association of Labour Economists (EALE), Oslo, 2007.

O'Connell, P. J. and Byrne, D. (2009), The determinants and effects of training at work: bringing the workplace back in, ESRI Working Paper, 289. Dublin: ESRI.

OECD (2005). Ageing and employment policies, Paris: OECD.

OECD (2008). Employment Outlook. Paris, OECD.

Oosterbeek, H. (1998). Unravelling supply and demand factors in work-related training, *Oxford Economic Papers* 50, 266–283

Riphahn, R. T. and Trübswetter, P. (2006), Population aging and continued education, IZA Discussion Paper 2415. Bonn: IZA

Sousa-Poza, A. and Henneberger, F. (2003), The determinants and wage effects of course-related training of elderly workers in Switzerland, Research Institute for Labour Economics and Labour Law Discussion Paper, 94. St. Gallen: University of St. Gallen.

Topel, R. H. and Ward, M. P. (1992), Job mobility and the careers of young men, *Quarterly Journal of Economics* 107, 439–479.

Veum, J. R. (1996), Training and job mobility among young workers in the United States, *Journal of Population Economics* 10, 219–233.

Appendix to Chapter 2

Table 2.4. Variable list for theoretical model

Variable name	***Variable description***
T	Binary training participation
R	Total net rents from training
ΔP	Total increase in productivity due to training
ΔP_t	Increase in productivity due to training in period t after training
C	Costs of training
a	Age when training takes place
r	Retirement age
t	Period after training
γ	Learning parameter
β	Depreciation rate

Table 2.5. Descriptive statistics of explanatory variables

Age categories	**Mean**
Age category 15–19 (dummy)	0.0513
Age category 20–24 (dummy)	0.1582
Age category 25–29 (dummy)	0.1946
Age category 30–34 (dummy)	0.1968
Age category 35–39 (dummy)	0.1968
Age category 40–44 (dummy)	0.1545
Age category 45–54 (dummy)	0.0478
Age polynomials	
Age at end of year (years) [standard deviation: 7.85; min.: 15; max.: 54]	31.9303
Age squared / 100	10.8124
Age cubed / 1000	38.4732
Age quartic / 10000	142.5961
Schooling (reference: no degree)	
Low school degree («Hauptschule») (dummy)	0.7209
Higher school degree (at least «Realschule») (dummy)	0.0799
Apprenticeship (reference: no degree)	
Apprenticeship degree outside firm (dummy)	0.4803
Apprenticeship degree in firm (dummy)	0.2514

Note: Number of observations is 10,544 from 415 workers in a balanced panel design.

Chapter 3

Effects of Training on Employee Suggestions and Promotions: Evidence from Personnel Records*

Joint work with

Christian Pfeifer, Simon Janssen, Uschi Backes-Gellner

*This chapter is based on the publication Pfeifer, C., Janssen, S., Yang, P., and Backes-Gellner, U. (2013), Effects of Training on Employee Suggestions and Promotions: Evidence from Personnel Records, *Schmalenbach Business Review* 65, 270-287.

3.1 Introduction

Returns on human capital investments have received considerable attention in policy and research over recent decades (e.g., Bartel, 1995; Bishop, 1997; Bartel. 2000; Asplund, 2005; Frazis and Loewenstein, 2005). Next to schooling, human capital accumulation after entry into the labor market is considered key to economic performance at both the micro and the macro levels. However, researchers have some problems when studying the impact of employer-provided formal training on workers' productivity. These difficulties include the aggregation of heterogeneous training types across industries and firms, and also the lack of adequate variables to serve as proxies for productivity. For example, survey data for workers compare individuals across firms with different training programs and often use workers' wage increases as a proxy for productivity increases. On the one hand, although wages might indeed be good proxies for productivity in perfect labor markets, they are obviously not so useful in imperfect labor markets. On the other hand, survey data on firms comprise only information about aggregated productivity (e.g., sales), which allows a comparison between firms but not between workers. Moreover, survey data often suffer from imprecise or even false statements about wages, training, and other variables. To overcome some of these problems, researchers have recently used the personnel records of single firms (e.g., Bartel, Ichniowski, and Shaw, 2004; Shaw, 2009). Although personnel data sets are not representative and are only "insider econometric" case studies, they have the advantage of comparing workers in the same firm, job, and training, and comparing unbiased information about wages, productivity, and training.

Another potential problem when evaluating causal effects of training is that training participation is likely to be non-random. Thus, if participation depends on unobservable characteristics, a cross-section comparison between workers who do participate in training and workers who do not is likely to suffer from omitted variable or selection bias. Panel data that exploit within variances can help deal with this problem, because first differences or conventional fixed-effects estimators address the issue of unobserved heterogeneity. Panel data compares outcomes such as the wages or productivity of a specific worker before and after training. Several empirical studies have recently used longitudinal data to close the research gap, but most attempts still suffer from measurement and aggregation biases in survey data. Moreover, few data sets provide sufficiently long panels to be able to exploit the time dimension in more detail. But the length of training effects in particular is important if we are to gain an understanding of actual depreciation rates of human capital investments, which are largely unexplored.

In this paper, we evaluate the causal effects of training at the lowest micro level by using personnel records from one German company. The data allow us to follow 415 male blue-collar workers, who entered the company in the late 1970s, over most of their working life, i.e., for more than 20 years. In addition to information about participation in formal training courses, our data set provides unique information about employee suggestions that were of productive value for the firm. Although missing information about training costs make it impossible for us to calculate returns on investments (ROI) and the actual benefits and costs of implementing workers' suggestions, we think that there are two reasons the analysis of training effects on the probability that workers will make suggestions is nevertheless important. First, because employee suggestions have not been used previously to study training effects. Thus, they are an interesting alternative to the often used supervisors' performance ratings in personnel data, which might suffer from subjectivity bias. Second, because employee suggestions are important for firms to permanently improve the efficiency of their production processes. Although training and suggestion systems are often idiosyncratic to firms, the question of whether training increases the probability of making suggestions for productivity improvements is general in nature.

We further analyze training effects on promotions, which we define as upward movement from one wage group to another and are hence associated with a wage increase. Promotions are important from the point of view of both employer and employee. Employees benefit from promotions by monetary gains and higher reputation, and employers can use promotions to make efficient job assignments. On the one hand, training might increase individual productivity by teaching skills and knowledge that are important to fulfill tasks at higher job levels. On the other hand, training can serve as a screening device without increasing individual productivity, i.e., the firm learns about abilities and skills of workers and can promote the best fitting (i.e., the most productive) worker to the next job in the hierarchy. Moreover, promotions are also determined by institutional arrangements such as seniority rules in collective contracts. Therefore, the link between training and promotions might be distorted by factors other than productivity. However, we think that employee suggestions are a good proxy for productivity effects from training.

To estimate the causal effects of formal training on the likelihood of workers making suggestions and getting promotions, we use individual fixed-effects linear probability and logit models. Our fixed-effects approach helps to mitigate problems that stem from unobserved heterogeneity and non-random training participation. We further exploit the length of the panel by constructing four lagged training variables, which make it possible for us to analyze the length of training effects. Thus, we are able to identify whether the effects of training on productivity and promotions are short term or long term.

The main findings of our econometric case study are that past training participation has significant positive effects on present suggestion and promotion probabilities. Training has the greatest impact on suggestion and promotion probabilities in the year directly after participation. The further in the past the training participation is, the more the training effect decreases in size and significance. This finding emphasizes the importance of the provision of employer-provided training throughout each employee's working life, not just in the early years of employment.

This paper is structured as follows. In Section 2 we summarize previous empirical findings on the effects of employer-provided training. In Section 3 we present the personnel data set, provides descriptive statistics, and discuss the econometric framework. In Section 4 we give the estimation results. Section 5 concludes.

3.2 Literature review

Following the pioneering contributions by Becker (1962) and Mincer (1974), a substantial body of economics literature on investments in human capital has addressed the determinants[16] and outcomes of training. A reason for the continuously growing number of empirical studies on training outcomes is rooted in recent advancements in overcoming methodological challenges and new data when trying to identify a causal effect of training participation.

The problem with the methods used to evaluate training effects is based on the potential endogeneity of the training variable. One source of this endogeneity stems from the concern about selection bias. Firms expect training participation to be unevenly distributed across workers with different abilities. Firms are likely to select those workers for training for whom the expected returns are most favorable (Leuven and Oosterbeek, 2008). Endogeneity of the

[16] For literature reviews on the determinants of training participation see Becker (1993), Leuven and Oosterbeek (1999), Neumark and Washer (2001), Leuven (2004), and Metcalf (2004).

training variable might lead to omitted variable bias. If training represents one of many determinants of wages and productivity, then the training effects could be over- or underestimated (Barron et al., 1999). To correct for endogeneity, recent empirical training studies draw primarily on approaches such as a Heckman-type selection (Lynch, 1992; Veum, 1995), instrumental variables (Leuven and Oosterbeek, 2008), or fixed effects estimation (Booth, 1993; Barron et al., 1999).

Despite the improved methods now used to correct for endogeneity, data availability still represents a major problem for three main reasons. First, because few studies find instruments that might affect training, yet not impact the outcome variable (Leuven and Oosterbeek, 2004). Second, because most panel data sets are relatively short, either variation is low or training cases are rare (Dearden et al., 2006). Short panels also do not allow the researcher to make inferences about the length of training effects through the use of lagged variables (Frazis and Loewenstein, 2005). Third, despite increased efforts to find adequate measurements of training participation, few studies obtain distinct outcome variables that unambiguously denote promotions in hierarchy and productivity on the individual level (Bartel, 2000).

Most empirical studies on training outcomes have addressed the wage effects of training participation (for reviews see Bishop, 1997; Bartel, 2000; Asplund, 2005). However, the effects of training on workers' promotions in hierarchies and on productivity has not received as much attention. The main explanation for this lacuna is that wages, according to traditional human capital theory, are an adequate proxy for hierarchy and productivity. In perfect labor markets, wages are equal to the value of marginal products of workers, at least for general training (Becker, 1962). Accordingly, promotions serve as recognitions of workers' increased productivities (Frazis and Loewenstein, 2005). However, in imperfect labor markets, employers are able to pay employees below their marginal product (Acemoglu and Pischke, 1998). Increased wages from training participation would then fail to serve as the proxy for the enhanced productivity of workers. Since several empirical studies find significant variations of wages within job levels (Baker et al., 1994a, 1994b; Lazear and Oyer, 2013), a wage increase is not necessarily associated with more responsibility at work or a shift to higher job levels. Therefore, recent empirical literature on training outcomes emphasizes the need to distinguish between wages, promotions, and productivity (Asplund, 2005).

Recent studies on wage effects of employer-provided training in Germany use matching and difference-in-differences methods (Muehler et al., 2007) or quasi experimental approaches in which a random event prevents a worker from taking part in scheduled training (Fahr and Simons, 2008). Muehler et al. (2007) identify twofold training effects in the German Socioeconomic Panel when differentiating between general and specific training. Consistent with standard human capital theory, only general training increases wages significantly. Fahr and Simons (2008) find low wage effects of training participation in a German cross-sectional survey on further training. However, a limited period of time between training and wage information limits their investigation of wage effects to the short term. Other recent studies that deal with the endogeneity of training and wages are Gerfin (2004), who applies nonparametric matching and difference-in-differences methods on Swiss Labour Force Survey data; and Schøne (2004), who uses fixed effects regression methods on the Norwegian Survey of Organisation and Employees.

Frazis and Loewenstein (2005) use survey data from the National Longitudinal Study of Youth and the Employer Opportunity Pilot Project to evaluate the effect of training on subsequent promotions. Workers self-report promotions and indicate if they have received a promotion in hierarchy or whether their job responsibilities have increased. The authors estimate fixed-effects regressions and find positive effects for current and past training participation on promotion probabilities. However, surveys entail subjective responses of individuals, which are likely to be subject to measurement errors (Bartel, 1995). Furthermore, the training variable underlies significant heterogeneity, so despite the effort to enhance the informational value of training measures through the observation of hours spent on training spells, questions remain as to how adequate the aggregation of different training types is.

Krueger and Rouse (1998) examine the impact of workplace education programs for one blue-collar and one white-collar company. They limit training heterogeneity by observing one standardized type of training form. This form was partially governmentally financed and was undertaken at the local community college between 1991 and 1995. By estimating an ordered probit model, the authors find that compared to untrained workers, trained workers are much more likely to make job bids and to receive job upgrades. Yet, the results suffer from a relatively low number of observations and insufficient panel length. Instead of using econometric approaches to limit selection bias, Krueger and Rouse have to assume that selection is controlled for by sufficient information on observed characteristics.

Most empirical studies on training effects on productivity use industry data or matched employer–employee data (Bartel, 2000). This slowly growing branch of literature typically uses the standard Cobb-Douglas production function and observes firms over several years. Empirical studies at the plant level primarily uses survey data of firms in the United States (Black and Lynch, 1996, 2001), the UK (Dearden et al., 2006), Italy (Conti, 2005), Germany (Zwick, 2005), and Ireland (Barrett and O’Connell, 2001). In general, most of these studies find that the share of trained workers has a positive effect on labor productivity, which diminishes with the inclusion of human resource management characteristics. However, few empirical studies have investigated productivity effects of training participation at the individual worker level.

Pischke (2001) uses data from the German Socioeconomic Panel from 1986 to 1989. He observes detailed information on workers' participation in formal training programs. As a training outcome, Pischke uses workers' responses on benefits from training participation. He finds support for a positive effect of formal training on self-reported performances of workers and interprets this finding as increased productivity.

Bartel (1995) recommends the use of data from personnel records of a single firm (econometric case study) for three main reasons. First, because personnel records provide exact training time and type. Second, because the training of workers is done by the same firm, corresponding to more homogeneous training measures. Third, because workers' outcomes are more comparable if they work for the same firm. Bartel (1995) uses personnel records from a large manufacturing company for the period from 1986 to 1990. To determine the effect of training on productivity, she uses information on performance evaluations by supervisors. Formal training has a positive and significant effect on the performance evaluations of workers, from which Bartel concludes that formal training has a productivity-increasing effect. However, the short panel does not allow any inferences on the length of training effects, and supervisors' performance ratings might suffer from potential biases such as subjectivity. A recent study by Breuer and Kampkötter (2010) uses three years of personnel records from a German multinational company and fixed-effects methods. These authors find that training only has a positive effect on several performance-related outcomes in the same year that training participation takes place. But we note that the research design might suffer from the short panel length.

As a form of innovative behavior and knowledge sharing tool, training and employee suggestions in particular have received substantial interest in human resource management and personnel psychology. Innovative behavior after formal training roots in two approaches. First, according to human capital theory, formal training increases ability and productivity. In a comprehensive review on the benefits of training for individuals, Aguinis and Kraiger (2009) note several channels through which training might directly affect individual knowledge and performance. Second, following Ford (1996), Schmidt (2007), and Arnold and Davey (1996) employer- provided formal training might indirectly enhance innovative behavior and creative thinking ability through organizational commitment, capability belief, or workplace satisfaction. Barber's (2004) qualitative study of data on Indian mechanics emphasizes the direct effect of training on innovation and performance. In an exploratory study of 372 employees of one firm, Cabrera, Collins, and Salgado (2006) identify self-efficacy, openness, and experience as the driving determinants of knowledge sharing. Organizational commitment and rewards are important, but these qualities affect innovative outcomes to a much lower degree than human capital development.

Clearly, the potential endogeneity of the training variable demands sophisticated econometric methods if it is to determine the causal effects of training participation on distinct outcomes such as wages, promotions, and productivity. Although there are several approaches for estimating causal effects, data availability represents a major problem. Panels are usually short so that the variation of training and outcomes is low. Furthermore, few data sets offer precise information on training and outcome variables. The training variable in survey data is usually aggregated through heterogeneous training types across firms and industries. As training outcomes, most empirical papers use wages as their proxies for hierarchy or productivity, and those that actually observe hierarchy and productivity rely on either heterogeneous outcomes or subjective evaluations. We complement these studies by using an insider econometric approach with long balanced panel data for one firm. This panel comprises unique information about training and outcomes such as employee suggestions and promotions. The data set makes it possible for us to apply fixed-effects estimation techniques with lagged training variables to make inferences about the length of training effects.

3.3 Personnel data, variables, and econometric method

We analyze the personnel records of a large company from the energy sector in western Germany. The company is subject to a collective contract and has a works council. Due to data protection reasons we are not allowed to name the company or to give detailed information. The data that we obtain from the personnel records comprise yearly information about a subsample of 438 blue-collar workers in the company's mining business. These workers entered the firm in four subsequent cohorts from 1976 until 1979 and stayed in the company over the entire observation period up to 2002. The sample represents a share of about a quarter of all employees in the company's operation unit and 3.5% of the company's entire workforce.

For our analysis, we restrict the sample to German male blue-collar workers that do not have any missing values in the variables we use. This restriction reduces our sample by 5% to 415 different workers. Since we are interested in the long-term effects of training, we use four lags of training participation, so that the first four yearly observations of every worker are dropped from the estimation sample. Moreover, we drop all observations from the last observation year, 2002, from the estimation sample, because no promotion variable can be constructed for that year. The final sample contains 8,469 yearly observations of 415 different workers. The number of workers is n = 105 for the entry cohort 1976. The observations included in the estimation sample for entry cohort 1976 ranges from 1980 to

2001, which leads to a panel length in years of T = 22. For entry cohort 1977: n = 96, T = 21. For entry cohort 1978: n = 77, T = 20. For entry cohort 1979: n = 137, T = 19.

Nearly 20% of these blue-collar workers do not have any secondary school degree. About 72% have the lowest secondary school degree *(Volks-/Hauptschulabschluss),* and approximately 8% have at least successfully completed medium secondary school *(Realschule).* We also know that roughly 25% of these workers have no apprenticeship qualification, around 25% have completed their apprenticeship in the analyzed company, and the remaining 50% have performed their apprenticeship in other firms. We assign formal employer-provided training in the company to four different types: a short training course *(kurze Schulung)* (one or two days); a longer training course *(längere Schulung)* (up to several weeks); longer vocational re-training *(längere Umschulung)* (up to several weeks); and a longer academy of vocational training *(Berufsakademie)* (up to several weeks). We observe a total of 626 training cases. More than two thirds are short training courses; the other training types are almost equally distributed. Because of the small number of cases in most training types, we use a binary variable that takes the value of one if a worker participated in any kind of training and zero otherwise. To reduce heterogeneity in the training courses, we also analyze the effects of short training courses separately. Unfortunately, due to the company's restrictions, we do not have information about the direct and indirect costs of these training courses or about their actual contents. However, we do know that workers are paid during the training period and do not have to cover any direct costs. Thus, all costs are covered by the employer.

To evaluate the effects of formal training in the company, we use two outcome variables. The first outcome is a binary variable that indicates if a worker makes a suggestion. These suggestions have productive value for the firm and workers receive monetary rewards for them. Again, because of company restrictions, we do not know more about the value of the suggestions or the potential implementation costs. Since we analyze blue-collar workers in the mining business, we believe it is likely that most suggestions are about more efficient work arrangements. Formal training courses might teach new aspects in work arrangements or stimulate ideas about the current work arrangements, thus giving workers greater opportunities to make suggestions after they have received such training. We observe 356 suggestions by workers, which results in a yearly average of about 4%.

The second outcome variable we use to assess the training effects is a binary variable that indicates if a worker is promoted from one wage group in a given year (*t*) to a higher wage group in the subsequent year (*t* + 1). We obtain the underlying wage groups from the collective contract. Promotions are by definition associated with a significant wage increase, which might be explained by a productivity increase due to training. However, promotions are not determined exclusively by a worker's productivity, because training might serve as screening device for efficient job assignments. Further, promotions are also affected by institutional rules (e.g., seniority arrangements in collective contracts). Thus, promotions are less likely to be a good proxy for productivity than are employee suggestions. We observe 511 promotions in our sample, which results in a yearly average of about 6%.

We describe our econometric framework in equation (1). In principle, we estimate the impact of lagged training participation *T* of worker *i* on his outcomes *Y* in year *t*, which are worker suggestions and promotions. We also include a set of time-variant control variables *X* (age in years, squared age divided by 100, and wage groups as continuous variable) that might also serve as proxies for otherwise unobserved variables that vary over time, and which help us to identify the causal training effects. We denote time fixed effects as λ_t, worker fixed effects as ν_i, and use ε_{it} as the usual error term. We denote the parameters we estimate β and δ. We present descriptive statistics of the variables in Table 3.1.

$$Y_{it} = \beta_1 T_{i,t-1} + \beta_2 T_{i,t-2} + \beta_3 T_{i,t-3} + \beta_4 T_{i,t-4} + \delta X_{it} + \lambda_t + \nu_i + \varepsilon_{it} \qquad (1)$$

Table 3.1. Descriptive Statistics

	Mean	**Std. dev.**	**Min.**	**Max.**
Suggestion in *t* (dummy)	0.0420	0.2007	0	1
Promotion in *t* (dummy)	0.0603	0.2381	0	1
Training in *t* – 1 (dummy)	0.0661	0.2485	0	1
Training in *t* – 2 (dummy)	0.0653	0.2471	0	1
Training in *t* – 3 (dummy)	0.0634	0.2437	0	1
Training in *t* – 4 (dummy)	0.0582	0.2342	0	1
Short training in *t* – 1 (dummy)	0.0433	0.2036	0	1
Short training in *t* – 2 (dummy)	0.0413	0.1991	0	1
Short training in *t* – 3 (dummy)	0.0367	0.1881	0	1
Short training in *t* – 4 (dummy)	0.0314	0.1744	0	1
Age in *t* (years)	33.4290	6.5271	19	53
Age squared / 100	11.6010	4.4034	3.61	28.09
Wage group in *t*	7.0461	2.7482	2	19

Notes: Total number of worker-year observations is 8,469 for 415 blue-collar workers.

The coefficients of interest are the *βs*, which are the effects of formal employer-provided training on the probability that a worker makes a suggestion or gets promoted. Using the lags of training participation has the advantage of estimating the correct causal direction, because past training participation has to affect current outcomes. Moreover, by comparing the *βs* we can infer the length of training effects. The inclusion of time and worker fixed effects reduces the efficiency of the estimates but makes it more likely that estimates of the βs are consistent, because omitted variable biases are reduced. Since worker fixed effects are jointly significant in all estimated specifications, and since Hausman tests reject the null hypothesis of no systematic differences with random effects estimates, we choose to use only fixed-effects models. Because of potential problems in fixed-effects probit and logit models, such as a sample size reduction due to the loss of individuals without any within variation in the outcome variables and the incidental parameter problem (Heckman, 1981), we prefer to estimate fixed-effects linear probability models (LPM) using ordinary least squares (OLS). According to Angrist (2001), if the main objective is to estimate causal effects and not structural parameters, then linear models can be appropriate even for limited dependent variables. Nevertheless, we also apply a fixed-effects (conditional) logit model as a robustness check.

To provide consistent effects for the βs, the $T_{i,t-1}$ to $T_{i,t-4}$ must be strictly exogenously conditional on our variables in X_{it}, the time fixed effects λ_t, and the unobserved worker fixed effects v_i, i.e., $E(\varepsilon_{it} \mid X_{it}, T_{i,t-1}, T_{i,t-2}, T_{i,t-3}, T_{i,t-4}, \lambda_t, v_i) = 0$ must hold for all t. Hence, $T_{i,t-1}$ to $T_{i,t-4}$ must be uncorrelated not only with ε_{it} but also with $\varepsilon_{i,t-1}$ and $\varepsilon_{i,t+1}$. In our case, we might argue that the firm selects a worker for training because the worker made a particularly good suggestion in the previous period, which signals his ability to the employer. If this were the case, then $T_{i,t-1}$ to $T_{i,t-4}$ should be correlated with $\varepsilon_{i,t+1}$. Consequently our estimates of β would not be consistent, because the assumption of strict exogeneity would be violated.

Fortunately, Wooldridge (2002, 285) proposes a simple procedure to test for the violation of strict exogeneity in fixed-effects models. The idea of the test is to include the lead of the explanatory variable in the regression model and to estimate the model with a fixed-effects approach. If the lead of the explanatory variable shows a significant correlation with the outcome variable, then this correlation would indicate that the respective explanatory variable is correlated with the error term $\varepsilon_{i,t+1}$. In our application, we perform the test by incorporating $T_{i,t+1}$ into regression equation (1) and by testing the null hypothesis that the coefficient of $T_{i,t+1}$ is zero. If we have to reject the null hypothesis, then the test indicates that $T_{i,t-1}$ to $T_{i,t+4}$ are correlated with $\varepsilon_{i,t+1}$ and that the assumption of strict exogeneity is likely to be violated. We present the results of the F-tests for strict exogeneity along with our regression results in the following section.

3.4 Estimation results

We present our estimation results for the probability of employee suggestions in Table 3.2. As a reference model, we estimate the first specification by using a pooled linear probability model (LPM) without fixed effects. We estimate the next four specifications by using worker fixed-effects linear regressions. Specification 2 includes only the first lagged training participation variable and no time fixed effects (year dummies). The predicted probability that the average worker without training will make a suggestion is about 4%, and for an average worker who has received training during the last year, the probability is about 6.6% The absolute marginal effect of 2.6 percentage points is statistically significant (p = 0.011) and economically important (the relative marginal effect is 2.6/4 = 65%). When we compare these results with the reference model in Specification 1, it appears that the estimated effect of training in the pooled LPM model without worker fixed effects suffers from only a small non-significant upward bias. Specification 3 includes additional time fixed effects, which are jointly significant in an F-test. The estimated training effect is only slightly reduced to 2.4 percentage points. Specification 4 includes the complete four lags of training participation and no time fixed effects. Specification 5 also includes the time fixed effects. We see that the marginal effect of the first lag is slightly reduced to 2.4 and 2.2 percentage points, but is still highly significant. The other three lags, i.e., training participation at least two years ago, have no significant effects on the suggestion probability.

At the bottom of the table we present the F values for Wooldridge's test of strict exogeneity, which are small and not statistically significant. Hence, we cannot reject the null hypothesis that the coefficients of $T_{i,t+1}$ are zero in any fixed-effects regression, so that the assumption of strict exogeneity does not seem to be violated in this application.

Table 3.2. Effects of Training on Employee Suggestions

	(1) LPM	**(2) LPM**	**(3) LPM**	**(4) LPM**	**(5) LPM**	**(6) Logit**
Training in $t-1$	0.0276**	0.0260**	0.0238**	0.0240**	0.0221**	0.4000*
	(0.0111)	(0.0102)	(0.0102)	(0.0105)	(0.0104)	(0.2310)
Training in $t-2$				0.0098	0.0113	0.2354
				(0.0098)	(0.0097)	(0.2448)
Training in $t-3$				–0.0036	0.0002	–0.0245
				(0.0086)	(0.0086)	(0.2750)
Training in $t-4$				–0.0107	–0.0036	–0.1174
				(0.0080)	(0.0079)	(0.3034)
Age	0.0222***	0.0221***	0.0080**	0.0221***	0.0080**	0.2367
	(0.0024)	(0.0025)	(0.0033)	(0.0025)	(0.0033)	(0.2137)
Age squared / 100	–0.0298***	–0.0282***	–0.0113**	–0.0282***	–0.0114**	–0.1575
	(0.0036)	(0.0037)	(0.0049)	(0.0037)	(0.0049)	(0.2777)
Wage group	–0.0001	–0.0007	–0.0012	–0.0006	–0.0012	–0.0182
	(0.0007)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0734)
Year fixed effects	No	No	Yes	No	Yes	Yes
Worker fixed effects	No	Yes	Yes	Yes	Yes	Yes
R^2	0.0115	0.1778	0.1888	0.1781	0.1889	
F value (complete model)	43.68***	43.74***	961***	2533***	8.57***	
F value (strict exogeneity)		1.27	0.19	1.24	0.23	
Pseudo R^2 (McFadden)						0.1596
Chi^2 value						
(complete model)						25599***
Number of observations	8469	8469	8469	8469	8469	2979
Number of workers	415	415	415	415	415	146

Notes: Mean yearly suggestion probability for an average worker without training is approximately 4%. Coefficients of linear probability models for specifications (1) to (5) and fixed effects (conditional) logit model for specification (6). Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Column 6 in Table 3.2 includes a robustness check for the method and sample. We estimate a fixed-effects (conditional) logit model for the complete specification (all lags of training, time fixed effects) on a subsample of workers who actually make a suggestion during the observation period. The estimated coefficients support the findings from the linear estimates, that only the first training lag has a significant effect. A noteworthy result of the estimates in Table 3.2 is the inverted u-shape effect of age on the suggestion probability, which has its maximum around the ages 35 to 40 years. If suggestions are related to productivity, then this finding is consistent with the concave age-productivity profiles known from other studies.

Table 3.3 elaborates the estimation results for the probability that a worker gets promoted, which is associated with a significant wage increase. Specification 1 is again the linear probability model without fixed effects as a reference model. Specification 2 (first lag, worker fixed effects, no time fixed effects) shows an absolute marginal effect of 7.7 percentage points due to training in the last year, which is highly significant. An average worker without training has a predicted promotion probability of 5.5%, but an average worker with training has a predicted promotion probability of 13.2%. When we compare these results with the reference model in specification one, the estimated effect of training on promotions in the pooled LPM model without worker fixed effects seems to suffer from a significant upward bias of more than two percentage points or about a quarter, due to unobserved worker heterogeneity. If we include time fixed effects in Specification 3, then the estimated training effect, 8.25 percentage points, is even larger than in Specification 2. Specifications 4 and 5 include all four lags of training participation. The estimated effects for the first training lag do not change significantly. Furthermore, the effect of the second lag is not significant, but the effects of the third and fourth lags are again significant. However, we note that the differences between the effects of the second, third, and fourth lags are not statistically significant from each other at conventional levels. The third lag has a marginal effect of about four percentage points and the fourth lag of about three percentage points. If we compare these effects with the effect of the first lag, we see that they have only half the size. Again, we present the *F* values for Wooldridge's test of strict exogeneity at the bottom of the table. These values do not indicate a violation of the assumption of strict exogeneity. Column 6 in Table 3.3 also includes a fixed-effects (conditional) logit model for the complete specification (all lags of training and time fixed effects), which we estimate on a subsample of workers who have actually been promoted during the observation period. The estimated coefficients support the findings from the linear regressions. Further, we find in all fixed-effects specifications that workers at higher wage groups are less likely to be promoted.

One might argue that suggestions and promotions are related to each other. For example, supervisors might be more likely to choose for promotion a worker who has recently made a suggestion. Therefore, we repeat the linear estimates for the complete specification (all lags of training, worker, and time fixed effects) with additional control variables that include four lags of promotions in the suggestion regression and vice versa. Because these variables have no significant effects and the results in Tables 3.2 and 3.3 are essentially unchanged, we present the estimation results of this robustness check only in the Appendix (see Table 3.5).

Table 3.3. Effects of Training on Promotions

	(1) LPM	**(2) LPM**	**(3) LPM**	**(4) LPM**	**(5) LPM**	**(6) Logit**
Training in $t-1$	0.1005***	0.0774***	0.0825***	0.0783***	0.0830***	0.9977***
	(0.0155)	(0.0148)	(0.0148)	(0.0150)	(0.0150)	(0.1535)
Training in $t-2$				0.0124	0.0165	0.2420
				(0.0133)	(0.0133)	(0.1805)
Training in $t-3$				0.0390***	0.0415***	0.6637***
				(0.0133)	(0.0133)	(0.1746)
Training in $t-4$				0.0318**	0.0298**	0.4640**
				(0.0132)	(0.0133)	(0.1881)
Age	–0.0101**	0.0009	0.0012	0.0019	–0.0001	–0.0640
	(0.0041)	(0.0042)	(0.0054)	(0.0042)	(0.0054)	(0.1031)
Age squared / 100	0.0106*	0.0020	0.0044	0.0004	0.0061	0.2671*
	(0.0059)	(0.0061)	(0.0076)	(0.0061)	(0.0076)	(0.1552)
Wage group	0.0025**	–0.0383***	–0.0381***	–0.0393***	–0.0392***	–0.4447***
	(0.0010)	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0395)
Year fixed effects	No	No	Yes	No	Yes	Yes
Worker fixed effects	No	Yes	Yes	Yes	Yes	Yes
R^2	0.0168	0.1023	0.1109	0.1051	0.1139	
F value (complete model)	22.37***	4759***	11.78***	2942***	11.01***	
F value (strict exogeneity)		1.24	0.70	0.67	0.30	
pseudo R^2 (McFadden)						0.1229
Chi^2 value (complete model)						326.53***
Number of observations	8469	8469	8469	8469	8469	5757
Number of workers	415	415	415	415	415	281

Notes: Mean yearly promotion probability for an average worker without training is approximately 5.5%. Coefficients of linear probability models for specifications (1) to (5) and fixed effects (conditional) logit model for specification (6). Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3.4. Effects of Short Training

	Suggestion		Promotion	
	(1) LPM	**(2) Logit**	**(3) LPM**	**(4) Logit**
Short training in $t-1$	0.0375***	0.5419**	0.0299**	0.5092**
	(0.0144)	(0.2556)	(0.0143)	(0.2409)
Short training in t- 2	0.0325**	0.5304*	–0.0066	–0.3523
	(0.0140)	(0.2722)	(0.0118)	(0.3074)
Short training in $t-3$	0.0051	–0.0009	0.0248*	0.4534*
	(0.0121)	(0.3292)	(0.0147)	(0.2694)
Short training in $t-4$	0.0131	0.2874	0.0160	0.2497
	(0.0122)	(0.3647)	(0.0163)	(0.2844)
Age	0.0077**	0.2196	0.0010	–0.0218
	(0.0033)	(0.2139)	(0.0054)	(0.1014)
Age squared / 100	–0.0112**	–0.1399	0.0045	0.1854
	(0.0049)	(0.2780)	(0.0076)	(0.1527)
Wage group	–0.0020	–0.0356	–0.0387***	–0.4161***
	(0.0013)	(0.0733)	(0.0031)	(0.0378)
Year fixed effects	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes
R^2	0.1904		0.1053	
F value (complete model)	8.73***		10.01***	
Pseudo R^2 (McFadden)		0.1625		0.1017
Chi^2 value (complete model)		260.68***		270.11***
Number of observations	8469	2979	8469	5757
Number of workers	415	146	415	281

Notes: Coefficients of fixed effects linear probability models for specifications (1) and (3) and fixed effects (conditional) logit model for specifications (2) and (4). Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To further reduce heterogeneity in the training variable, in our next step we concentrate on short training courses. Short training courses are one- or two-day courses and make up about two thirds of all training cases in the data. For suggestion and promotion probabilities, we estimate fixed-effects linear models for the complete sample and fixed- effects logit models for subsamples of workers who actually make a suggestion or are promoted during the observation period. We present the results in Table 3.4. In general, the results are consistent with our previous findings on aggregated training. But two notable differences arise. First, the effect of short training on suggestions is greater and significant for the last two years. Second, the effect of short training on promotions is smaller. These differences between short training and aggregated training might be explained by different course contents and aims. Short training courses are likely to be more concerned with improvements in the current work arrangements and less with teaching completely new skills (e.g., retraining), which might be important to obtain better-paid jobs in the firm’s hierarchy. Thus, career-oriented longer training courses might indeed be more attractive for younger workers. However, short training courses, which seem to have only short-term effects on productivity, are still important for older workers (skill updating, employability) and are justified from an economic perspective because shorter amortization periods of old workers should play a minor role if depreciation rates are that large.

3.5 Conclusion

In this paper, we use the unique personnel records of a German company to evaluate the effects of formal employer-provided training on employee suggestions and promotions. By following this “insider econometric approach”, we are able to address such issues as training course heterogeneity and unobserved worker heterogeneity. We find significant positive, but only short-term, effects of training on the probability that workers will make suggestions, which indicates a high depreciation rate in this dimension. Moreover, we find that training participation increases the promotion probability. Our results are generally consistent with the human capital argument that training increases workers' productivities, but also with reciprocal behavior of workers in the context of gift-exchange (Akerlof, 1982; Leuven et al., 2005), i.e., workers react with increased effort to make suggestions in exchange for the training given by the company. However, the short-term effect of training raises the question of whether depreciation rates are greater than previously assumed and ROIs smaller than often computed. If this were the case, then the often-stated argument that old workers receive no training due to short amortization periods would no longer be that convincing. Because we use only a sample of blue-collar workers in one single firm and qualitative information about employee suggestions and promotions in an econometric case study, we cannot give definitive answers to this question. But we hope for more studies to come that use long panels of personnel data.

References for Chapter 3

Acemoglu, D. and Pischke, J.-S. (1998), Why do firms train? Theory and evidence, *Quarterly Journal of Economics* 113, 79–119.

Aguinis, H. and Kraiger, K. (2009), Benefits of Training and Development for Individuals and Teams, Organizations, and Society, *Annual Review of Psychology* 60, 451–474.

Akerlof, G. A. (1982), Labor contracts as partial gift exchange, *Quarterly Journal of Economics* 97 (4), 543–569.

Angrist, J. D. (2001), Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice, *Journal of Business and Economic Statistics* 19 (1), 2–16.

Arnold, J. and Davey, K. M. (1999), Graduates' Work Experiences as Predictors of Organisational Commitment, Intention to Leave and Turnover, Which Experiences Really Matter?, *Applied Psychology* 48 (2), 211–238.

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

Baker, G., Gibbs, M., and Holmstrom, B. (1994a), The Internal Economics of the Firm: Evidence from Personnel Data, *Quarterly Journal of Economics* 109 (4), 881–920.

Baker, G., Gibbs, M., and Holmstrom, B. (1994b), The Wage Policy of a Firm, *Quarterly Journal of Economics* 109 (4), 921–956.

Barber, J. (2004), Skill upgrading within informal training: Lessons from the Indian auto mechanic, *International Journal of Training and Development* 8(2), 129–139.

Barrett, A. and O'Connell, P. J. (2001), Does Training Generally Work? The Return to In-Company Training, *Industrial and Labor Relations Review* 54 (3), 647–662.

Barron, J. M., Berger, M. C., and Black, D. A. (1999), Do Workers Pay for On-The-Job Training, *Journal of Human Resources* 34 (2), 235–252.

Bartel, A. P. (1995), Training, wage growth, and job performance: Evidence from a company database, *Journal of Labor Economics* 13, 401–425.

Bartel, A. P. (2000), Measuring the Employer's Return on Investments in Training: Evidence from Literature, *Industrial Relations* 39 (3), 502–524.

Bartel, A. P., Ichniowski, C., and Shaw, K. L. (2004), Using “insider econometrics” to study productivity, *American Economic Review* 94 (2), 217–223.

Becker, G. S. (1962), Investments in human capital: a theoretical analysis, *Journal of Political Economy* 7 (5), 9–49.

Becker, G. S. (1993), Human Capital: A Theoretical and Empirical Analysis with a Special Reference to Education, *University Chicago Press*, Chicago.

Bishop, J. H. (1997), What we know about employer-provided training: a review of the literature, *Research in Labor Economics* 16, 19–87.

Black, S. E. and Lynch, L. M. (1996), Human–Capital Investments and Productivity, *American Economic Review*, Paper and Proceedings 86 (2), 263–267.

Black, S. E. and Lynch, L. M. (2001), How to compete: The impact of workplace Practices and Information technology on Productivity, *Review of Economics and Statistics* 83 (3), 434–445.

Booth, A. L. (1993), Private Sector Training and Graduate Earnings, *Review of Economics and Statistics* 75 (1), 164–170.

Breuer, K., and Kampkötter, P. (2010), The effects of intra-firm training on earnings and job performance – evidence from a large German company, Paper presented at EALE/SOLE Meeting 2010.

Cabrera, A., Collins, W. C., and Salgado, J. F. (2006), Determinants of individual engagement in knowledge sharing, *Internal Journal of Human Resource Management* 17 (2), 245–264.

Conti, G. (2005), Training, productivity and wages in Italy, *Labour Economics* 12 (4), 557–576.

Dearden, L., Reed, H., and van Reenen, J. (2006), The Impact of Training on Productivity and Wages: Evidence from British Panel Data, *Oxford Bulletin of Economics and Statistics* 68 (4), 397–421.

Fahr, R., and Simons, S. (2008), Returns to company training: Evidence from a new approach using quasi experimental data. Mimeo.

Ford, C. M. (1996), A Theory of Individual Creative Action in Multiple Social Domains, *Academy of Management Review* 21 (4), 1112–1142.

Frazis, H. and Loewensrein, M. A. (2005), Reexamining the returns to training: Functional form, magnitude, and interpretation, *Journal of Human Resources* 40 (2), 453–476.

Gerfin, M. (2004), Work-related training and wages: An Empirical analysis for male workers in Switzerland. IZA Discussion Paper No. 1078.

Heckman, J. J. (1981), The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process, in Charles F. Manski and Daniel McFadden (eds.), *Structural analysis of discrete data with econometric applications*, MIT Press, Cambridge, 179–195.

Krueger, A. B. and Rouse, C. E. (1998), The Effect of Workplace Education on Earnings, Turnover, and Job Performance, *Journal of Labor Economics* 16 (1), 61–94.

Lazear, E. P. and Oyer, P. (2013), Personnel Economics, Forthcoming in R. Gibbons and J. Roberts (eds.), Handbook of Organizational Economics, Princeton University Press, Princeton, 479–519.

Leuven, E. (2004), A Review of the Wage Returns to Private Sector Training, Paper presented at the EC_OECD Seminar in Human Capital and Labour Market Performance: Evidence and Policy Challenges. Brussels.

Leuven, E. and Oosterbeek, H. (1999), The demand and supply of work-related training: Evidence from four countries, *Research in Labor Economics* 18, 303–330.

Leuven, E. and Oosterbeek, H. (2004), Evaluating the Effect of Tax Deductions on Training, *Journal of Labor Economics* 22(2), 461–488.

Leuven, E. and Oosterbeek, H. (2008), An Alternative Approach to Estimate the Wage Returns to Private Sector Training, *Journal of Applied Econometrics* 23, 423–434.

Leuven, E., Oosterbeek, H., Sloof, R., and van Klaveren, C. (2008), Worker Reciprocity and Employer Investment in Training, *Economica* 72, 137–149.

Lynch, L. M. (1992), Private-Sector Training and the Earnings of Young Workers, *American Economic Review* 82 (1), 299–312.

Metcalf, D. (2004), The impact of the national minimum wage on the pay distribution, employment and training, *Economic Journal* 114(494), C84–C86.

Mincer, J. A. (1974), Schooling, Experience, and Earnings, Columbia University Press, New York.

Muehler, G., Beckmann, M., and Schauenberg, B. (2007), The returns to continuous training in Germany: New Evidence from propensity score matching estimators, *Review of Managerial Science* 1, 209–235.

Neumark, D. and Wascher, W. (2001), Minimum Wages and Training Revisited, *Journal of Labor Economics* 19 (3), 563–595.

Pischke, J.-S. (2001), Continuous Training in Germany, *Journal of Population Economics* 14 (3), 523–548.

Schmidt, S. W. (2007), The relationship between satisfaction with workplace training and overall job satisfaction, *Human Resource Development Quarterly* 18 (4), 481–498.

Schøne, P. (2004), Why is the return to training so high?, *Labour* 18, 363–378.

Shaw, K. (2009), Insider econometrics: a roadmap with stops along the way, *Labour Economics* 16 (6), 607–617.

Veum, J. R. (1995), Sources of Training and Their Impact on Wages, *Industrial and Labor Relations Review* 48 (4), 812–826.

Wooldridge, J. M. (2002), Econometric analysis of cross section and panel data, MIT Press, Cambridge.

Zwick, T. (2005), Continuing Vocational Training Forms and Establishment Productivity in Germany, *German Economic Review* 6 (2), 155–184.

Appendix

Table 3.5. Trainings Effects when Controlling for Promotion and Suggestion

	(1) Suggestion	**(2) Promotion**
Training in *t* – 1	0.0221**	0.0826***
	(0.0107)	(0.0150)
Training in *t* – 2	0.0102	0.0161
	(0.0101)	(0.0133)
Training in *t* – 3	0.0010	0.0410***
	(0.0088)	(0.0133)
Training in *t* – 4	–0.0037	0.0298**
	(0.0082)	(0.0133)
Age	0.0081**	–0.0003
	(0.0033)	(0.0053)
Age squared / 100	–0.0114**	0.0063
	(0.0049)	(0.0076)
Wage group	–0.0014	–0.0391***
	(0.0014)	(0.0030)
Promotion in *t* – 1	0.0009 (0.0092)	
Promotion in *t* – 2	0.0068 (0.0092)	
Promotion in *t* – 3	–0.0073 (0.0074)	
Promotion in *t* – 4	0.0029 (0.0076)	
Suggestion in *t* – 1		0.0129 (0.0154)
Suggestion in *t* – 2		0.0162 (0.0162)
Suggestion in *t* – 3		0.0041 (0.0158)
Suggestion in *t* – 4		–0.0118 (0.0145)
Year fixed effects	Yes	Yes
Worker fixed effects	Yes	Yes
*R*2	0.1891	0.1143
F value (complete model)	7.55***	9.64***
Number of observations	8469	8469
Number of workers	415	415

Notes: Coefficients of fixed effects linear probability models. Robust standard errors in parentheses.

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

Chapter 4

The Impact of Music on Educational Attainment

4.1 Introduction

The formation and development of human capital is of central importance to society and has attracted significant interest in recent economic research. Often related to economic growth and economic prosperity, human capital development extends to the creation of equal opportunities within society (Becker, 1993). The economics literature has mainly evaluated common initiatives to increase the stock of human capital initiatives such as formal education and on-the-job training. Alternative approaches, especially those that do not at first glance appear related to human capital, have yet to be explored.

Music is one potential human capital-increasing activity that has received little attention in the economics literature. Its beneficial effects on cognition have long been suspected, dating to antiquity (Rainbow, 1998) when philosopher Plato praised music for its powerful role in human life and society in shaping sensibility and rationality (Pelosi, 2010). The association of music activity with math is commonly attributed to Pythagoras. However, the notion that music strongly relates to cognitive skills has already been expressed on ancient clay tablets in Egypt around 2000 BC (Pont, 2004).

There are many reasons to believe that musical activity, used below to denote playing an instrument or singing, potentially affects educational attainment[17]. A vast body of literature has shown that music relates positively with cognitive and non-cognitive skills. Studies using magnetic resonance data suggest the development of distinctive brain structures among professional musicians compared to non-musicians. Glaser and Schlaug (2003) find larger music-related brain structures that are relevant for motor and auditory skills for musicians[18].

[17] Alternative music literature addresses the effect of music listening and music theory on cognition related outcomes. See Rauscher (1993), Kaempfe et al. (2011), Garlin & Owen (2006), Hetland (2000), Vaughn (2000) and Edelsohn & Johnson (2003).

[18] Similar results are obtained by Pantev et al. (1998) and Hyde et al. (2009).

With respect to cognitive skills, Elbert et al. (1999) and Pantev et al. (1998) observe better developed corpus callosum (CC) - the body connecting the left and right cerebral hemispheres that is central to an individual's intelligence- for professional musicians[19]. In consensus, although heavily criticized for its causal interpretation[20], music affects music-related and cognition-related brain development.

More convincing with respect to causality is experimental literature on the impact of music on cognition in the short-term[21]. Schellenberg (2004) evaluates an experiment in which 144 six-year-old students were assigned to a one-year music treatment (keyboard or voice lessons), a different treatment (drama lessons) or no treatment. He finds that students who received the music treatment outperformed the other groups in standardized academic achievement. Gardiner et al. (1996) survey an experiment that randomly allocated 96 first graders into music classes and classes without music, finding that the former exhibited significant improvements in reading, mathematics, classroom attitudes, and behavior ratings after one additional year of music class.

Overall, few studies have related music to educational outcomes under the utilization of larger sample sizes. Morrison (1994) examines a representative cross-sectional dataset from the National Center for Educational Statistics of some 18,000 students. Schellenberg (2006) evaluates an experiment combined with a cross-sectional dataset of 300 students. However, both studies do not hold individual and family information constant. Southgate and Roscigno (2009) relate current music activity to math and reading scores using US datasets from the Early Childhood Longitudinal Program (ECLS) and the National Education Longitudinal

[19] Also see Schlaug et al. (1995), Lee et al. (2002) and Schlaug & Jaencke (2001).

[20] For recent debates see Schellenberg (2006; 2011).

[21] For a broad based meta-analyses of IQ, see Schellenberg (2001; 2006; 2011), Billhartz et al. (2000), or Johnson & Memmott (2006). For broad based meta-analyses of spatial tasks, see Rauscher et al. (1997) and Hetland (2000). Costa Giomi (1999) is one of few experiments not to find any significant effect.

Study (NELS) holding constant an indicator for socio-economic status. Their findings suggest a positive relationship between current music activity and math scores for the sample of children and none for the sample of adolescents.

Theoretically, music may affect education through a variety of channels. Schellenberg (2011) suggests that music increases cognitive and non-cognitive skills by increasing cognitive functioning by transferring music-related skills such as memorization and visualization to non music-related domains. Also, music might improve cognitive capacities by enhancing executive functioning. This argument is based on the assumption that the complexity of music pieces, interpretation and analysis increases musically-unrelated problem-solving and judgment. According to Schumacher (2009), playing music and performing music pieces may further improve self concepts. In turn, better judgment of own and other’s achievements could stimulate competitiveness and ambitions. Hille and Schupp (2013) point out that playing music may affect personality traits such as openness and from a time allocation perspective denote a way to spend time that positively affects IQ. Next to these direct channels, it is possible that indirect factors such as the association with other peers and social well being play a significant role. When music tuition is provided by schools only, increased IQ may be driven by better institutions.

Winner et al. (2013) puts forward two mechanisms through that music may affect school performance. First, music might directly affect school outcomes through the association of music and particular school subjects. For example, music is likely to improve verbal skills such as reading writing and foreign language by facilitating auditory skills. Second, when music education is considered a school-like activity it may improve school-relevant skills such as concentration, patience, memorization, and commitment.

This article is the first economics paper[22] to evaluate the effects of musical activity in childhood on educational attainment in adolescence. It is the primary study to estimate the long-term effects of music on cognition-related outcomes. To achieve this, this study exploits a unique dataset from the German Socio-Economic Panel that makes it possible to use new identification strategies to study the effects of music. This paper primarily considers regressions under the conditional independence assumption (CIA), whereby the effects of variables that potentially affect both musical activity and educational attainment, such as ability, parents, and income, are held constant. Further, the biographical dataset is used to create a two-period panel dataset to evaluate the within effect on an individual who receives music lessons between the ages of ten and 17, using individual fixed-effects (FE) regressions. Both identification strategies are tested with instrumental variables estimations.

The biographical Youth Questionnaire of the German Socio-Economic Panel features a comprehensive set of data at the individual and household levels. In the present paper, this study observes a sample of 17-year-olds who answered retrospective, current and prospective questions related to educational path, leisure activities, and household characteristics. The Youth Questionnaire uniquely includes information on the highest educational level attained by parents, household income, and relationships with parents.

With respect to educational attainment, this study analyzes the effects of music on track recommendations after elementary school and school track attended at age 17. The German middle school system consists of three types of middle schools that differ in quality and curriculum; thus attending a better middle school track is highly relevant for future labor market outcomes (Dustmann, 2004) and wages (Cooke, 2003). The impact on education is estimated for different music models that observe different indicators of musical activity:

[22] Surprisingly, no economic study has examined the impact of musical activity on cognitive skills, while the related sports economics literature enjoys significant popularity. See Lechner (2009) and Pfeifer & Cornelißen (2010) for recent publications in the economics of sports.

paid music lessons, the start of one's musical activities, the intensity of one's musical activities, and whether music is played alone or with others.

Using multivariate regression analysis under CIA and FE models, this study finds a positive effect of music on track recommendations and track at age 17. The ordinary least squares and ordered probit results under CIA are highly significant for every specification and music model examined. In particular, the coefficients from ordered probit regression indicate that musical activity increases the predicted probability of receiving the highest track recommendation by around 20 percent and in turn decreases the probability to receive a lower track recommendation by around 15 percent.

The results of the fixed-effects models are somewhat similar to the findings obtained under the conditional independence assumption. Noteworthy, the impact of music remains relatively stable even when only considering German states with binding track recommendations. Furthermore, robustness checks that include instrumental variables models to validate fixed-effects regressions support the previous results. When instrumenting the music variable with church affiliation measured by church visits or membership in the Protestant church music still has a positive and significant effect on educational attainment.

Musical activity affects educational attainment in the long-term. When distinguishing musicians by the date at which music instruction began, it becomes apparent that the impact of music on educational outcomes increases the earlier music lessons begin. The estimates for the effects of music lessons in early childhood are significantly larger than in all other model specifications. Moreover, the intensity of musical activity matters. The music effect is substantially larger when music is played on a daily basis than when it is played less frequently. Surprisingly, no difference in the music coefficient is found with respect to paid music lessons versus any kind of lessons.

The subsequent paper is structured as follows. The next section provides information on the data used in this study. Section 3 illustrates the identification strategy used to estimate the impact of music on educational outcomes. The regression analyses are presented in Section 4. The paper concludes with a brief summary and a discussion of the results.

4.2 Data

4.2.1 BIOAGE 17

The German Socio-Economic Panel (SOEP) provides users with a representative collection of biographical information of the entire German population. Since 2000, the Youth Questionnaire BIOAGE17 (henceforth "BIOAGE17") has collected responses to youth specific questions from a sample of adolescents aged 17 at the time of the initial interview. For sophisticated information on data collection and coding, see Frick and Samberg (2007). Uniquely, BIOAGE17 features a rich set of retrospective, current, and prospective questions related to educational path, leisure activities, and household characteristics, including information on past track recommendations, last school track, musical activity, type of music lessons undertaken, the intensity of musical activity, the beginning of musical activity, the mode of musical activity, gender, the number of siblings, town size, county of residence, and mobility during childhood. Furthermore, BIOAGE17 elicits various statements regarding current relationships with parents and parental involvement in school- and leisure-activities. A detailed description on how the data was cleaned and combined with other SOEP datasets is to be found in the data appendix.

BIOAGE17 includes several reports on music. Most important, a dummy variable on music ("music active") indicates whether the respondent is active in music. This very general indicator, however, refers simply to present engagement, without any reference to intensity or context. For greater precision, the adolescent respondents also report how often music is currently played and whether music is played alone or in company. With respect to intensity, the data allow us to distinguish adolescents who make music on a daily basis ("music often") from those who make music on a weekly/monthly basis ("music seldom"). Concerning the context of music, it is observable whether the respondent makes music alone ("music alone") or with others ("music with").

Two retrospective questions about music lessons and the beginning of musical activity deepen the understanding of the relationship between adolescents and music. It is very likely that the quality of music lessons is determined by the professionalism of the teacher. To proxy for the quality of music lessons, a binary variable for paid music lessons ("music paid") indicates whether the respondent received paid music lessons during childhood. Uniquely, the biographical BIOAGE17 dataset allows for identification of the age when musical instruction began. The present paper distinguishes among adolescents who began to receive music lessons before age six ("music early childhood"), between ages six and ten ("music childhood"), and after age ten ("music youth").

Cognition-related data in the BIOAGE17 dataset only exist with respect to scholastic achievement, with the main outcome variables being track recommendation and last school track attended at age 17[23]. The relevance of German middle school tracks in the determination of wages and careers has been well documented in economics research[24]. A detailed description of the German middle school system is presented in the data appendix. Alternative educational outcome measures such as individual marks in subjects featured too many missing observations. Most importantly, this paper estimates the effect of music

[23] Some responds have reported individual test scores in different school subjects. However, these scores are missing for a large proportion of the sample that it seemed advisable not to use these for the analysis.

[24] See Dustmann (2004) and Cooke (2003).

activity on the track recommendation after elementary school. German students receive a track recommendation by teachers after elementary school that can be considered some type of third party evaluation of child's educational performance. This recommendation, in most federal states, is based on the average marks but also entail a proportion of subjective assessment of the teacher as in how good a child will fare in a certain middle school track. Track recommendation ("track rec") are coded "1" for lower, "2" for intermediate, and "3" for higher recommendations. Next to the track recommendations, this paper further reports the effect of music variables on the last track at age 17. The data on track at age 17 does not have any missing observation but it is for also another reason, why this outcome variable matters. Using the track at age 17, it is possible to study the long term effect of music on educational attainment. Both outcome variables, track recommendations and track at age 17, may very well differ because of non-binding track recommendations in certain German states where parents do not have to follow these recommendations and second chances of the German middle school system Dustmann et al. (2012). Both aspects are discussed in the Appendix. The last track at age 17 is coded 0 for dropouts, "1" for lower track, "2" for intermediate, and "3" for higher track.

Personal characteristics may affect music participation and scholastic achievement. Thus, it is beneficial to include this information in the regression analyses. This study controls for gender, number of siblings, and immigration status. Dummy variables are used to indicate a male respondent ("male") and the existence of siblings ("sibling"), while a binary migration background variable ("immigration") broadly indicates any type of migration background. Access to music may also differ with respect to town size and mobility during childhood. A discrete variable on town size during childhood ("town size") differentiates between cities, ranging from small to very large. Mobility is indicated by a dummy variable ("move child"), which indicates whether an adolescent moved during childhood. In addition, the adolescents report in the German state in which they currently reside. This information was used to generate dummy variables for each German state to control for state specific effects in the regressions. Unfortunately, the dataset does not include any information on the children's ability. Parental education is the sole available proxy for the ability of respondents.

The importance of parents in the music decision process and school performance is undeniable. Most importantly, this study is able to control for the highest level of education completed by both parents, using binary variables for degrees higher than a vocational degrees ("father educmed"; "mother educmed") and at least a high school diploma and above ("father educhigh"; "mother educhigh"). Household resources are proxied by a continuous variable reporting the logarithm of net monthly income in the year 2000 ("loghhincome 2000"). Note that the data on the logarithm of net monthly goes along with a sizable reduction of observations. The empirical analysis with information on household resources is therefore only used for robustness checks. To eliminate year specific effects, this study is only capable to include binary year variables of the interview year to proxy information on years. The sample that this study uses has been interviewed between 2001 and 2009.

BIOAGE17 includes information on the relationships between adolescents and their parents and parental influence. Although these statements are likely to be subjective, it may still be useful to control for them. These reports are either negatively formulated in that the influence of parents is exerted through pressure and arguments or positively formulated in that the influence of parents comes from support and engagement with teachers and schools. The amount of pressure felt from parents ("parents push") is coded 1, for no pressure, up to 4, if adolescents feel heavily pressured by their parents. A dummy variable indicates whether the respondent has problems with his or her parents because of school issues ("parents problems school"). Three dummy variables report statements regarding whether adolescents feel their parents are helpful in school affairs (“parents help”), regularly attend school nights (“parents school nights”), and consult teachers (“parents teachers”).

Lastly, two variables with respect to religion from BIOAGE 17 are used to instrument the potentially endogenous music variable. Church affiliation is a dummy variable that is one if the respondents visit church at least on a seldom basis. Protestant membership is a binary variable that is one if the respondent is a member of the Protestant church. Note that both variables obtain a substantial amount of missing observations and are therefore only used for robustness checks.

4.2.2 Summary statistics

Table 4.1 presents summary statistics of the main variables.

Table 4.1. Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs.**
Music variables					
Music active	0.233	0.423	0	1	3222
Music paid	0.171	0.377	0	1	3222
Music often	0.090	0.286	0	1	3222
Music seldom	0.135	0.341	0	1	3222
Music early childhood	0.056	0.230	0	1	3222
Music childhood	0.097	0.300	0	1	3222
Music youth	0.079	0.270	0	1	3222
Music alone	0.093	0.291	0	1	3222
Music with	0.140	0.347	0	1	3222
School variables					
Track school	2.233	0.716	0	3	3222
Track no	0.007	0.082	0	1	3222
Track low	0.125	0.331	0	1	3222
Track med	0.397	0.489	0	1	3222

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs.**
Track high	0.471	0.499	0	1	3222
Track rec	2.283	0.758	1	3	2595
Track rec low	0.186	0.389	0	1	2595
Track rec med	0.345	0.476	0	1	2595
Track rec high	0.469	0.499	0	1	2595
Personal characteristics					
Male	0.507	0.500	0	1	3222
Sibling	0.700	0.827	0	1	3222
Immigration	0.215	0.411	0	1	3220
Town size	1.284	1.136	0	3	3202
Move child	0.889	0.314	0	1	3222
Parental variables					
Mother higher educated	0.224	0.417	0	1	3222
Mother med. Educated	0.585	0.492	0	1	3222
Father higher educated	0.317	0.465	0	1	3222
Father med. Educated	0.557	0.497	0	1	3222
Loghhincome 2000	7.825	0.447	5.19	9.46	3222
Parents push	1.981	0.756	1	3	3215
Parents help	0.773	0.419	0	1	3214
Parents problems school	0.550	0.497	0	1	3222
Parents school nights	0.222	0.416	0	1	3222
Instrumental variables					
Church affiliation	0.518	0.500	0	1	1149
Protestant	0.344	0.475	0	1	1124

Music, generally learned through paid lessons during childhood, is frequently played alone. Some 23 percent of the sample is active in music. A considerable number of respondents make music on a weekly or monthly basis. Fewer respondents, nine percent of the total, are involved in music on a daily basis. Notably, approximately nine percent of the sample makes music only by themselves, while 14 percent of the sample also makes music with others. Approximately 17 percent of all interviewees received paid music lessons during childhood.

A reasonable number of respondents began music in early childhood or youth. Most musicians, approximately ten percent of the sample, began participating in music during childhood between the ages of six and 12.

Nearly half of the 17-year-olds last attended a higher middle school track, while slightly fewer, but approximately 40 percent of the sample, were in intermediate middle school tracks at age 17. A minority of the sample completed a lower middle school track or dropped out of school. Surprisingly, these shares are very similar to the track recommendations, despite some 20 percent of the sample missing observations for recommendations. Nearly 19 percent of the sample received lower track recommendations after elementary school. The majority obtained intermediate track recommendations - approximately 35 percent of the sample - or higher track recommendations - approximately 47 percent.

Additionally, Table 4.1 reports summary statistics for the variables at the individual level. The sample includes equal shares of male and female respondents, and most respondents have siblings. Some 22 percent of the respondents have some type of migration background. The adolescents were evenly distributed across city sizes during childhood, and nearly 90 percent of the sample moved between cities during childhood.

In addition to variables at the individual level, Table 4.1 presents summary statistics of the variables related to the household, in particular to parents. The majority of parents, some 59 percent of mothers and 56 percent of fathers, completed a vocational degree. Higher educational degrees were obtained by 22 percent of mothers and 31 percent of fathers. The mean value for parental pressure of approximately 1.9 indicates that most adolescents are neither left to their own devices nor under heavy pressure from parents. Many respondents, approximately 77 percent, stated that they receive help from parents with school work. The number of adolescents experiencing problems with their parents because of school issues is nearly identical to the number without any school-related problems with parents. Only approximately 22 percent of adolescents reported that parents regularly attended school nights.

4.3 Identification and estimation

The estimated effect of music is causal if the music treatment is randomly assigned. However, there are two main reasons why music participation, in a non-experimental world, is not random.

First, ability may affect the decision to learn and continuously play music (Ceci and Williams, 1997). As mentioned in the literature review section, MRI reports of larger musically relevant brain regions and better developed CCs in the neuroscience literature stress the existence of a positive correlation between music and brain structures. If innate ability has a positive effect on music participation and school performance, the estimated effect of music may be biased (Monaghan et al., 1998; Petrill et al., 2004; Plomin et al., 1997). Then music does not have an effect on cognition, yet innate ability (Neisser et al., 1996; Wechsler, 1991) or educational achievement (Ceci and Williams, 1997) determines participation in music activity. This is known as ability bias in the economics literature.

Second, it is reasonable to assume that peers and family background affect the decision to learn and play music (Hall, 2005). Musical activity entails substantial costs for initial music lessons and instruments. Therefore, household resources may be a limiting factor in children’s access to and participation in music (Gottfredson, 2002; Gottfredson and Lapan 2002; Schmidt and Hunter, 1998). Moreover, as documented in psychological studies, supportive and better-educated parents (Sergeant and Thatcher, 1974) are most likely to enroll their children in music lessons and supervise home practice (Sichivitsa, 2007, Zdzinski, 1996). If these household characteristics affect school outcomes, the music effect is also likely to be biased.

To adjust for endogeneity caused by ability and selection, this paper relies on regressions based on the CIA (Section 4.1.) and FE models (Section 4.2.). Under CIA assumption, also called selection on observables, regression coefficients may be interpreted as causal if all relevant covariates are known and controlled for, as music treatment is then as good as randomly assigned (Angrist and Pischke, 2009). However, it is arguable that some relevant confounding variables may remain unobserved. For this reason, this study also considers FE regressions to adjust for unobserved but fixed omitted variables (Wooldridge, 2002). Although the BIOAGE17 data are originally cross-sectional form, this paper is able to exploit the time dimension of the biographical survey to identify two periods, before and after musical experience begins, to estimate the within effect of music on an individual who receives music lessons between ages ten and 17. Finally, the robustness of the results based on the CIA and FE models is tested using instrumental variable regressions (section 4.3).

4.3.1 Conditional independence assumption

This study uses ordinary least square regressions (OLS) to estimate the effect of various music variables on the outcome variables under CIA. Each regression is repeated with an increasing number of control variables. The final specification reads:

$$y_i = ß_0 + ß_1 m_i + ß_2 X_i + ß_3 F_i + ß_4 inc_i + ß_5 c_i + ß_6 year_i + \varepsilon_i \tag{1}$$

where y_i denotes the outcome variable, either track recommendation or school track at age 17 of an individual i. Variations in the outcome variables are explained by music m_i, a variable that either indicates general musical activity, the intensity of musical activity, whether music is played alone or also with others, whether music lessons are paid for, or the time when music lessons began. X_i holds constant the effects of gender, siblings, immigration status, size of town during childhood, and mobility during childhood. F_i consists of dummy variables for the highest educational levels attained by both parents as well as variables that control for individual statements about parents, pressure received from parents, arguments with parents, parental assistance with school matters and parental attendance at school nights. inc_i is the logarithm of monthly net household income and c_i is a dummy variable each German federal state of residence. $year_i$ consists of several dummy variables that hold constant the effect of the interview year the best available proxy for the missing year information and ε_i is the error term.

This final specification that includes all relevant variables is estimated with an ordered probit model as well. Mainly, this serves two main purposes. First, the application of ordered probit regressions is to indicate whether the previous OLS results suffer from bias related to the linearity assumption. Second, using ordered probit additionally yields predicted probabilities for each track separately and therefore indicates particular effects of music for the probability to receive a certain recommendation or to be in a certain track at age 17. The ordered probit model under CIA is then:

$$y_i^* = ß_1 m_i + ß_2 X_i + ß_3 F_i + ß_4 inc_i + ß_5 c_i + ß_6 year_i + \varepsilon_i \quad = x_i'ß + \varepsilon_i \qquad (2)$$

with the distinct probabilities for recommendations:

$$Pr(y_i = 1 \mid x_i = x) = \Phi(\mu_1 - x'ß) \qquad (3)$$

$$Pr(y_i = 2 \mid x_i = x) = \Phi(\mu_2 - x'ß) - \Phi(\mu_1 - x'ß) \qquad (4)$$

$$Pr(y_i = 3 \mid x_i = x) = 1 - \Phi(\mu_2 - x'ß) \tag{5}$$

where $\Phi(.)$ denotes the standard normal distribution function.

4.3.2 Regression using individual fixed-effects

Despite inclusion of the above-noted individual and family variables in the OLS regressions under the CIA presented above, it is likely that relevant information remains unobserved. If the included control variables do not suffice to ensure random assignment to music, the results will suffer from omitted variable bias and will not be causal. For this reason, this paper employs panel data models, in particular FE models, to cope with potential omitted variables. Assuming that the unobserved effects are time invariant, FE analysis may well yield consistent results regarding the music effect (Wooldridge, 2002). To apply panel data models, it is necessary to exploit the time dimension of BIOAGE17, initially a cross-sectional, biographical dataset. Figure 4.1 illustrates how the biographical data were used to generate a two-period panel.

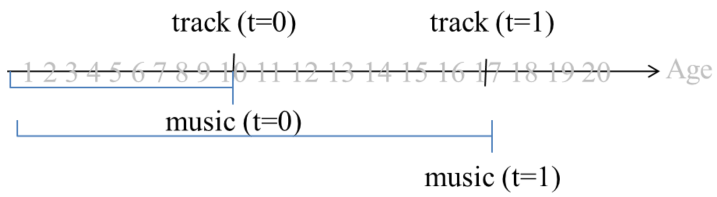


Figure 4.1. Panel data

Following Borjas' (2000) to exploit time information in biographical data to generate a two-period panel dataset, this study uses BIOAGE17 information to identify a track and a music variable for two periods: period t=0, when the respondents were ten years old, and period t=1, when the respondents were 17-years-old. The track variable in period t=0 indicates track recommendation, while the track variable in period t=1 indicates school track at age 17, obtained from BIOAGE17. The underlying assumption is that both track variables capture the same thing. Note that the coding, fortunately, is identical, as no dropouts reported their track recommendations; thus, the track information for both periods is coded 1 for “lower track”, 2 for “intermediate track” and 3 for “higher track”. To generate music variables for both periods, this study employs the BIOAGE17 data on when music lessons began. Music in period 0 is a dummy variable that takes a value of 1 if a respondent received music lessons before age ten. Music in t=1 is a dummy variable that takes a value of 1 if the respondent received music lessons before age 17. Unfortunately, this paper cannot identify any respondents who quit music after a certain number of years. According to BIOAGE17, all adolescents receiving music lessons in childhood were still active in music at the time of the interview at age 17. Using the two period panel dataset, this paper focuses primarily on the results of the FE model:

$$y_{it} = ß_0 + ß_1 m_{it} + v_i + \varepsilon_{it} \tag{6}$$

where y_{it} denotes the school track y of individual i during period t. Under the assumption of time invariant omitted variables, $ß_1$ denotes the causal effect of music m on an individual i at time t. Finally, v_i is a time invariant fixed-effect, and ε_{it} is an error term. This estimation if repeated for a subsample of respondents from German federal states with binding track recommendations after elementary school. Mainly, these results from the subsample serve to show that the estimated difference for whole sample is not solely driven by parents who do not follow the track recommendations. An overview on the German regulation with respect to binding and non-binding track recommendations is to be found in the Appendix.

As in the section before, ordered probit regressions are used to validate the results from OLS. Unfortunately, there is no feasible implementation of individual fixed effects in ordered probit models to estimate predicted probabilities. However, it is possible to instead use the Chamberlain Mundlak device to exploit the panel data (Wooldridge, 2002) with:

$$y_{it}^{*} = x_{it}'ß + z_i'\delta + \varepsilon_{it} \tag{7}$$

where z_i contains the time averages of all time varying variables with the distinct probabilities to be in a certain track after music exposure:

$$Pr(y_{it} = 1 \mid x_{it} = x, z_i = z) = \Phi(\mu_1 - x'ß - z'\delta) \tag{8}$$

$$Pr(y_{it} = 2 \mid x_{it} = x, z_i = z) = \Phi(\mu_2 - x'ß - z'\delta) - \Phi(\mu_1 - x'ß - z'\delta) \tag{9}$$

$$Pr(y_{it} = 3 \mid x_{it} = x, z_i = z) = 1 - \Phi(\mu_2 - x'ß - z'\delta) \tag{10}$$

4.3.3 Instrumental variables

A further concern of the previous estimation strategies addresses the limitations of individual fixed-effects estimators in uncovering a true causal effect of treatment in the presence of attenuation bias. Attenuation bias exists due to measurement errors in treatment variables that would otherwise tend to remain constant over time; then changes interpreted as effects of the treatment may actually be mostly noise (Angrist and Pischke, 2009). Although the two-period fixed-effects model used in this paper is unlikely to suffer from attenuation bias, this potential bias is examined.

A possible solution to attenuation bias caused by measurement error in fixed-effects models is instrumental variables (Angrist and Pischke, 2009). To validate the results of the fixed-effects regression, this study uses an instrumental variables (IV) model that instruments music activity with religious affiliation and being a member in the protestant. It is plausible to assume that religious affiliation is positively correlated with musical activity. The majority of churches, especially the Protestant churches, have choirs that offer affordable individual singing lessons and joint choir lessons. Moreover, many churches incorporate and support instrument lessons in the church to attract children. Thus, it is likely that individuals closely affiliated with churches have greater access to music than those unaffiliated with churches. Moreover, church affiliation is unlikely to be correlated with the error term. In is debatable whether church affiliation or membership in the Protestant church as an instrumental variable is more likely to satisfy the exclusion restriction compared to parental education or state of residence. Sander (1992) finds that most of the effect of religion on education is denotable to family characteristics, in particular parental education. However, there is a large amount of literature that observes a significant relationship between religion and education[25]. For this reason, the IV results are to be considered robustness checks only.

4.4 Results

The analysis of music-related school performance begins with a presentation of regression results on the effects of music during childhood on educational attainment, where the latter, under the CIA (Section 4.1), is proxied by track recommendation after elementary school. Then, this paper reports estimates of the impact of music on educational attainment in adolescence, proxied by school track at age 17 (Section 4.2). The particular effect of music on the probability to receive a certain track recommendation or to be in a certain track is documented in the following (Section 4.3). Then, this paper provides regression results of FE models (Section 4.4). As regressions under the CIA are likely to leave relevant variables

[25] Darnell & Darren (1997) as well as Lehrer (1999) demonstrate that fundamentalists belief and conservative Protestant affiliation decrease educational attainment. Sandner & Krautmann (1995) observe a positive effect of Catholic schooling on high school graduation.

unobserved, these methods are used to alleviate omitted variable bias. Finally, this paper shows the results of the robustness checks using IV (Section 4.5).

4.4.1 Educational attainment in childhood

Table 4.2 presents the coefficients and standard errors from the various OLS regressions under the CIA. In each regression, the dependent variable is the track recommendation of a teacher after elementary school, where 1 is the lowest track and 3 is the highest. Each column of Table 4.2 corresponds to one of the five models, where each uses different music-related information. Whereas the first model (column 1) includes a regression coefficient for general music activity, the subsequent models are more specific in that they relate the outcome variable to paid music lessons (column 2), the intensity of musical activity (column 3), the beginning of music lessons (column 4), and potential peer effects (column 5). The rows correspond with particular specifications of each model, each with successively more control variables. The final specification has a smaller number of observations, as several data points on household income are missing. Note that this paper only reports the coefficients of the music variables.

Overall, music has a positive and large effect on track recommendations. This effect is significant for all indicators of music involvement and for each specification in Table 4.2. General music activity (column 1) increases track recommendations by 0.418. This coefficient is fairly stable when individual control variables are included but declines with the inclusion of variables for parental characteristics and parental education. Nevertheless, the music effect remains positive and significant around 0.265 and does not substantially change when variables for household income are added in the final specification. The standard errors associated with all the music variables are small.

Table 4.2. Track recommendation and music

Music	1 **Active**	2 **Paid**	3 **Intensity**		4 **Begin**		5 **Peers**	
			Often	**Seldom**	**0-5 years**	**6-10 years**	**Alone**	**With**
No controls								
Coeff.	0.418***	0.471***	0.472***	0.415***	0.556***	0.473***	0.408***	0.426***
(s.e.)	(0.031)	(0.032)	(0.042)	(0.037)	(0.045)	(0.040)	(0.043)	(0.037)
N	2,567	2,567	2,567		2,567		2,567	
R-squared	0.058	0.059	0.063		0.063		0.058	
Individual controls								
Coeff.	0.395***	0.437***	0.461***	0.383***	0.506***	0.455***	0.371***	0.412***
(s.e.)	(0.030)	(0.032)	(0.041)	(0.037)	(0.046)	(0.039)	(0.043)	(0.036)
N	2,567	2,567	2,567		2,567		2,567	
R-squared	0.098	0.097	0.102		0.101		0.098	
Individual and parental controls								
Coeff.	0.265***	0.277***	0.339***	0.244***	0.325***	0.324***	0.208***	0.304***
(s.e.)	(0.030)	(0.033)	(0.039)	(0.037)	(0.049)	(0.038)	(0.043)	(0.035)
N	2,567	2,567	2,567		2,567		2,567	
R-squared	0.223	0.223	0.227		0.226		0.224	
Individual, parental, income, year and state controls								
Coeff.	0.260***	0.250***	0.317***	0.249***	0.327***	0.292***	0.2145***	0.288***
(s.e.)	(0.035)	(0.040)	(0.050)	(0.044)	(0.065)	(0.048)	(0.052)	(0.042)
N	2,101	2,101	2,101		2,101		2,101	
R-squared	0.257	0.244	0.260		0.257		0.258	

Note: This table reports various OLS estimates of the relationship between track recommendation (1-3) and music indicators. Each of the five regressions was estimated using four specifications with increasing numbers of confounding variables. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In comparison to general music activity, the coefficients for paid music lessons (column 2) are larger in the first two specifications. Paid music lessons, as an indicator of teaching quality, are associated with higher recommendations absent any control variables. However, when including control variables for parents and income, the coefficients for general musical involvement and paid music lessons are somewhat similar. Unsurprisingly, the intensity of music participation is important (column 3). The effect of music on track recommendation is larger when music is played on a daily basis than when it is played less frequently. The difference in the music effect between musicians who practice often and those who practice seldom grows as more control variables are added.

Most importantly, this study finds a relationship between track recommendations and the beginning of music lessons, indicated as occurring either before six years of age ("early childhood") or from six to ten years of age ("childhood") (column 4). Music in early childhood has the largest effect on track recommendation; children who begin practicing music before age 6 obtain higher recommendations than non-musicians, an increase of 0.566, an estimate that is far larger than that for any other music variable and remains larger even when holding individual and family variables constant. A similar yet somewhat smaller trend is also observable for the coefficients of music in childhood. Note that the reference groups for both dummy variables are either respondents without any musical experience or respondents whose musical instruction began after ten years of age. Evidently, the beginning of music lessons matters. The earlier that instruction begins, the larger are the potential effects on school outcomes in the future. With respect to peer-effects, this paper finds a small advantage if music is played with others (column 5). The correlations between music and track recommendation are somewhat larger for those who do not exclusively play music by themselves.

Generally, Table 4.2 emphasizes the positive effect of music on school outcomes, an effect that is highly significant in each model and specification. As expected, parental education and involvement are highly important control variables, as after inclusion of these variables, the regression coefficients decrease significantly (rows 3-4). Surprisingly, the music coefficient does not change when individual, income, or state controls are added.

4.4.2 Educational attainment in adolescence

To relate music to educational outcomes in the long-term, Table 4.3 reports the coefficients and standard errors of various OLS regressions, under the CIA, of school track at age 17 on music. For each regression, the dependent variable is the ordinal information on school track at age 17, where 0 indicates dropout, 1 indicates at least lower track, 2 indicates at least intermediate track, and 3 indicates the highest track. As noted in the previous section, some shifting between tracks may have occurred during middle school, and some parents may have sent their children to a higher track school, despite a lower track recommendation, if those recommendations were non-binding. For these reasons, there are deviations between track recommendation and school track at age 17. Each column of Table 4.3 refers to one of the five models that use different music-related information. Whereas the first model (column 1) considers general musical activity, the subsequent models are more specific in that they relate the outcome variable to paid music lessons (column 2), the intensity of musical activity (column 3), when music lessons begin (column 4), and potential peer effects (column 5). Each row constitutes another specification of each model, each with increasing numbers of control variables. The final specification has fewer observations, as several data points on household income are missing. For brevity, this paper only reports the coefficients of the music variables.

Table 4.3. School track at age 17 and music

Music	1 **Active**	2 **Paid**	3 **Intensity**		4 **Begin**			5 **Peers**	
			Often	**Seldom**	**0-5 years**	**6-10 years**	**11-17 years**	**Alone**	**With**
No controls									
Coeff.	0.417***	0.487***	0.455***	0.412***	0.612***	0.453***	0.245***	0.448***	0.396***
(s.e.)	(0.026)	(0.026)	(0.036)	(0.031)	(0.032)	(0.034)	(0.044)	(0.037)	(0.031)
N	3,183	3,183	3,183			3,183		3,183	
R-squared	0.061	0.066	0.063			0.071		0.062	
Individual controls									
Coeff.	0.407***	0.467***	0.451***	0.396***	0.579***	0.445***	0.249***	0.424***	0.396***
(s.e.)	(0.025)	(0.025)	(0.035)	(0.031)	(0.033)	(0.034)	(0.043)	(0.036)	(0.030)
N	3,183	3,183	3,183			3,183		3,183	
R-squared	0.095	0.096	0.097			0.103		0.095	
Individual and parental controls									
Coeff.	0.264***	0.295***	0.312***	0.245***	0.373***	0.288***	0.169***	0.246***	0.275***
(s.e.)	(0.024)	(0.026)	(0.032)	(0.030)	(0.034)	(0.032)	(0.041)	(0.035)	(0.029)
N	3,183	3,183	3,183			3,183		3,183	
R-squared	0.239	0.238	0.240			0.241		0.239	
Individual, parental, income, year and state controls									
Coeff.	0.254***	0.282***	0.293***	0.250***	0.378***	0.260***	0.178***	0.230***	0.268***
(s.e.)	(0.031)	(0.035)	(0.044)	(0.038)	(0.058)	(0.043)	(0.047)	(0.045)	(0.036)
N	2,648	2,648	2,648			2,648		2,648	
R-squared	0.257	0.257	0.260			0.260		0.249	

Note: This table reports various OLS estimates of the relationship between school track at age 17 (0-3) and music indicators. Each of the five regressions was estimated using four specifications with increasing numbers of confounding variables. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Overall, a positive and significant impact of music on school track at age 17 is observed in all models and specifications (columns1-5). Surprisingly, although a significant number of observations on track recommendation are missing initially, the results for track recommendation and school track at age 17 are somewhat similar. Musical activity increases the last track attended by 0.417, absent any control variables. When all relevant control variables are included, the music effect decreases to 0.264, yet it remains highly significant. Here again, the regression results change significantly when parental information is considered.

The effect of music on the last track attended is larger for paid music lessons and daily practice. Note that for school track at age 17, the model that estimates the correlation with the beginning of music instruction now includes a dummy variable for musicians who began music lessons between ten and 17 years of age, where the reference group consists only of those not involved in music (column 4). With this specification, it becomes apparent that the long-term effects of music are larger than the short-term effects. Early music lessons increase the final track attended by 0.612 (versus 0.245 when lessons begin later), suggesting that the effect is larger, the earlier that instruction starts. With respect to peer effects, the relationship is slightly stronger for music within a group framework. Generally, the specifications point to the importance, when evaluating the relationship between music and educational outcomes, of the confounding variables of parent’s education and involvement. Individual controls, household income and state controls do not change the coefficients significantly.

4.4.3 Particular music effects

As discussed in the previous chapter, the OLS results are not intuitively interpreted. Also, under consideration of the linearity assumption, the OLS results are tested using an empirical model that potentially fits better to the ordinal outcome variable. Table 4.4 reports predicted probabilities to receive a certain track recommendation or to be in a certain track at age 17 in dependence of several music indicators from an ordered probit model. Note that all regressions hold constant the effect of individual characteristics, family background, income year and state.

Table 4.4. Results from Ordered Probit Model

	Track recommendation			School track at age 17			
	Low	**Intermediate**	**High**	**No**	**Low**	**Intermediate**	**High**
Prob. at means	18.6%	34,5%	46.9%	0.6%	12.5%	39.7%	47.1%
(1) Music activity							
Coeff.	-11.4%***	-4.6%***	16.0%***	-1.0%***	-9.5%***	-7.3%***	17.8%***
(s.e.)	(0.015)	(0.006)	(0.020)	(0.002)	(0.011)	(0.008)	(0.019)
N		2101				2648	
Log likelihood		-1880.70				-2308.17	
(2) Paid music lessons							
Coeff.	-11.7%***	-4.7%***	16.4%***	-1.1%***	-11.6%***	-8.7%***	21.4%***
(s.e.)	(0.017)	(0.007)	(0.024)	(0.003)	(0.013)	(0.010)	(0.023)
N		2101				2648	
Log likelihood		-1887.31				-2307.1552	
(3) Intensity of play							
Coeff. (often)	-14.7%***	-5.9%***	20.6%***	-1.1%***	-11.7%***	-8.9%***	21.6%***
(s.e.)	(0.023)	(0.009)	(0.031)	(0.003)	(0.017)	(0.012)	(0.029)
Coeff. (seldom)	-10.6%***	-4.2%***	14.8%***	-0.9%***	-9.2%***	-6.9%***	17.1%***
(s.e.)	(0.019)	(0.008)	(0.025)	(0.003)	(0.014)	(0.010)	(0.024)
N		2101				2648	
Log likelihood		-1875.84				-2303.17	
(4) Begin of music lessons							
Coeff. (0-5 years)	-16.1%***	-6.4%***	22.5%***	-1.8%***	-17.7%***	-13.3%***	32.8%***
(s.e.)	(0.031)	(0.012)	(0.042)	(0.004)	(0.025)	(0.018)	(0.044)
Coeff. (6-10 years)	-14.0%***	-5.5%***	19.5%***	-1.0%***	-10.0%***	-7.5%***	18.5%***
(s.e.)	(0.022)	(0.009)	(0.029)	(0.002)	(0.016)	(0.012)	(0.028)
Coeff. (11-17 years)	-	-	-	-0.6%***	-5.7%***	-4.3%***	11.6%***
(s.e.)	-	-	-	(0.001)	(0.016)	(0.012)	(0.029)
N		2101				2648	
Log likelihood		-1877.06				-2298.48	
(5) Peers							
Coeff. (alone)	-9.1%***	-3.7%***	12.8%***	-0.9%***	-9.0%***	-6.9%***	16.8%***
(s.e.)	(0.022)	(0.009)	(0.031)	(0.002)	(0.017)	(0.012)	(0.030)
Coeff. (with)	-12.8%***	-5.1%***	17.9%***	-1.0%***	-9.7%***	-7.5%***	18.3%***
(s.e.)	(0.018)	(0.008)	(0.025)	(0.002)	(0.013)	(0.010)	(0.023)
N		2101				2648	
Log likelihood		-1879.81				-2308.09	

Note: Probabilities denote average partial effects. Each of the five regressions is calculated for the probability to receive a certain track recommendation (Low, Intermediate, High) or for the probability to be in a certain track (No, Low, Intermediate, High) at age 17. Robust standard errors are in parentheses. Individual, parental, income, year and state fixed-effects are held constant for all regressions. *** p<0.01, ** p<0.05, * p<0.1

Generally, the results from the ordered probit models support the existence of a significant and positive effect of music in the long-term. Music, observed through different music indicators, increases the probability to receive a higher track recommendation by around 20 percent. At the same time, music activity decreases the probability to receive a lower track recommendation tremendously. This effect is even slightly larger for the impact of the music variables on the school track at age 17. As in the OLS results, the music effect is larger when music lessons started early and if music was played more intensively.

4.4.4 Fixed-effects results

As noted above, it is possible that the previously reported regressions under the CIA leave key variables unobserved. Thus, the results of the first three sections should not be interpreted as causal, as selection into the music treatment is not random if important variables are not controlled for. Therefore, this study also considers FE models to address possible omitted variable bias. Table 4.5 reports estimations of FE and ordered probit models with Chamberlain-Mundlak device of the effects of music on track attendance. Next to the whole sample, this paper further reports the fixed effects results for the subsample of those students who live in German federal states that have binding track recommendations (column). These results are to test that the estimated difference between last track at age 17 and track recommendation is not only the result of parents allocating their children into higher track recommendation than initially recommended in the second specification.

Table 4.5. Panel data results (FE)

Balanced Panel	**(1)**	**(2)**	**(3)**			**(4)**		
	Track-all	**Track-binding**	**Track-all**			**Track-binding**		
			Low	**Intermediate**	**High**	**Low**	**Intermediate**	**High**
Music	0.175***	0.161***	-6.8%***	-5.0%***	11.8%***	-6.4%**	4.8%**	11.2%**
(s.e.)	(0.043)	(0.059)	(0.025)	(0.018)	(0.043)	(0.035)	(0.026)	(0.060)
Constant	1.317***	1.322***	-					
(s.e.)	(0.011)	(0.012)	-					
Observations	5180	2910	5180			2910		
Method	FE	FE	Chamberlain-Mundlak device			Chamberlain-Mundlak device		
R-squared	0.0611	0.0535	-			-		
R-squared (within)	0.0068	0.0053	-			-		
R-squared (between)	0.0773	0.0675	-			-		
F-test/Wald	16.52	7.51	-			-		
Log likelihood	-	-	-4949.839			-2796.58		
Number of id	2,590	1455	2590			1405		
rho	0.678	0.682	-			-		

Note: This table reports various OLS-individual fixed effects estimates of the relationship between school track at age 17 (0-2) and music activity. The results are analog to a first difference estimator. A Hausman test that was used to compare the fixed effect model to a random effects model rejects the null hypothesis of unsystematic differences in the coefficients, indicating that the individual fixed effects are likely to be correlated with the music variable and thus potential omitted variable bias in the random effects model. The fixed effect model was estimated using Chamberlain-Mundlak (CM) device, yielding the same coefficient and almost identical R-squared. Note that FE controls for both portion of the unobserved individual levels variance, the one correlated with the average music and the one uncorrelated with that average, whereas CM only controls for the portion of the variance correlated with the average of music. The CM device indicated the significance of the unobserved individual effects. Specification 1 reports the effect for the whole sample. In the second specification, this paper only observes students from German federal states that give out binding track recommendations after elementary schools. The third specification reports probabilities that denote average partial effects estimated through an ordered probit model that uses CM device for the whole sample. In specification 4 the results include only those respondents from German states with binding track recommendations. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Most notably, column (1) of Table 4.5 reports a positive and significant effect of music on school track under FE. When holding individual effects constant, music participation increases school track by 0.175. Although this is substantially smaller than the regression results under CIA (always slightly above 0.2, after including all control variables), it remains highly significant. This effect remains relatively stable when only considering students who live in a German federal state that have binding track recommendation, an indication that the deviance of the track recommendation and the track at age 17 are not the outcome of parents who allocate their children into higher school tracks than recommended.

4.4.5 Further robustness checks

To validate the fixed-effects results this paper additionally estimates IV models of the music effect by instrumenting the potentially endogenous music variable by church affiliation. Table 4.6 reports IV results, instrumenting musical activity through church affiliation and being a member of the Protestant church, on track recommendation and last track attended. As mentioned before, these estimates are by far to be interpreted as causal, since previous literature on religion and educational outcomes do not exclude a significant impact of religious affiliation on education. However, they are still noteworthy to be understood as robustness checks for the estimates of the previous sections.

Table 4.6. Robustness checks using IV

	A: Church affiliation			
	Track recommendation		**School track at age 17**	
Model	OLS	IV	OLS	IV
Instrument	-	Church affiliation (0/1)	-	Church affiliation (0/1)
Coef.	0.219***	-	0.225***	-
(s.e.)	(0.048)	-	(0.042)	-
ATT (music acitivity)	-	1.101***	-	0.985***
(s.e.)	-	(0.322)	-	(0.243)
First stage (church affiliation)	-	1.609***	-	0.178***
(s.e.)	-	(0.029)	-	(0.027)
Individual controls	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman Test Endogeneity (p-value)	-	0.001	-	0.004
Test of joint significance (p-value)	-	0.000	-	0.000
Partial R squared-first stage	-	0.029	-	0.087
F-test-first stage	-	29.83	-	43.75
N	961	961	1139	1139

	B: Protestant			
	Track recommendation		**School track at age 17**	
Model	OLS	IV	OLS	IV
Instrument	-	Protestant (0/1)	-	Protestant (0/1)
Coef.	0.292***	-	0.329***	-
(s.e.)	(0.057)	-	(0.049)	-
ATT (music acitivity)	-	0.943	-	1.678**
(s.e.)	-	(0.615)	-	(0.077)
First stage (Protestant)	-	0.083***	-	0.067***
(s.e.)	-	(0.030)	-	(0.025)
Individual controls	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman Test Endogeneity (p-value)	-	0.214	-	0.002
Test of joint significance (p-value)	-	0.216	-	0.002
Partial R squared-first stage	-	0.030	-	0.068
F-test-first stage	-	7.51	-	6.82
N	876	876	1120	1120

Note: This table reports OLS and IV estimates for the effect of music on track recommendations (1-3) and school track at age 17 (0-3). The OLS regressions were estimated with the same sample that has been used for the IV model and includes the same set of control variables. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The estimates of the linear IV models of Table 4.6 suggest a highly significant and positive music effect. However, the results of the IV estimation must be interpreted with caution. In particular, it should be noted that the sample size is significantly smaller under this specification, as many respondents did not respond to church-related questions. Additionally, it is likely that there is considerable endogeneity, as church affiliation is hardly random. Nevertheless, it is noteworthy that the positive and significant effect of music remains, even the music variable is instrumented through church affiliation. Solely for track recommendation, when instrumenting musical activity through Protestant membership, the music coefficient is not significant. However, the size of the coefficient is almost the same as the music coefficient instrumented through church affiliation. It is very likely that the lost significance is caused by the low observations. This assumption is supported by the positive and significant estimate of the music effect on school track at age 17 instrumented through Protestant membership using a slightly larger sample.

4.5 Conclusion

A vast amount of literature in consensus agrees on a positive correlation between music and cognition-related outcomes. Using MRI scans, medical reports have found increased CC for professional musicians, leaving unclear if these improvements were induced by music or if innate ability itself has determined music participation. More convincingly with respect to the causal interpretation, experiments have identified a significant and positive impact of music on cognition in the short-term by randomizing music treatment and evaluating intelligence before and after music. However, no evidence exists with respect to the impact of music in the long-term.

This paper is the first economics study of the effects of music on cognition-related outcomes. In particular, it evaluates the effect of musical activity in childhood on educational outcomes in childhood and adolescence, complementing the existing literature by producing the first long-term estimates of the music effect. Using multivariate regressions under CIA and FE models it estimates the impact of music while holding constant the effects of individual and family characteristics. In this respect, this paper benefits substantially from the rich biographical datasets of the SOEP, which offer, in addition to individual information, comprehensive information about parents and children's interactions with parents. Furthermore, this paper uses distinctive music indicators to distinguish the specific effects of paid music lessons, the intensity of musical activity, the age at which musical instruction begins, and whether music is made alone or with others.

In general, musical activity has a positive and significant effect on track recommendation and track at age 17, an effect that is highly significant under every specification and music model. The music coefficient of music on last track attended is approximately the same magnitude as that of track recommendation, after holding all relevant variables constant. When estimating the particular probability to receive a certain track recommendation or to be in a certain middle school type at age 17, this study finds that music increases the probability of higher track recommendation or being in a higher track by around 20 percent for all music indicators. This effect is highly significant. In turn, music activity decreases the probability to receive a lower recommendation or to be in the lowest middle school track by around 15 percent.

The results of the FE models are somewhat similar to those under CIA, suggesting a positive and highly significant effect of music on educational attainment. Furthermore, the size of the FE coefficients does not differ substantially from those obtained under CIA. Even when only considering the difference between track recommendation and last track attended for the sample of those students in German federal states with binding recommendations, the results do not differ significantly.

Alternative approaches such as instrumental variable regressions confirm the robustness of the music coefficient. When instrumenting the potentially endogenous music variable by church affiliation or membership in the Protestant church the music impact remains significant and positive. This is surprising, considering the much smaller sample being used.

Interpreting the coefficients from the music models in detail, the results of this paper suggest that paid music lessons, when used to indicate high quality instruction, do not have substantially larger effects on school track than ordinary lessons. After considering all relevant control variables, the impact of paid music lessons and general musical activity are somewhat similar. The music effect, unsurprisingly, is affected by the intensity of musical activity. Music has a much larger effect on school track if it is played on daily basis than if it is played less frequently. Most importantly, the beginning of music instruction matters. The earlier music lessons begin, the stronger is the music effect. Thus, the impact on school track is largest if music lessons begin in early childhood. The identification of the channels through which the music effect is able to extend to the long-term, even excelling the longer music is played, is an interesting research question for studies to come. This paper is unable to identify a peer effect.

Notably, successive inclusion in the regressions of control variables demonstrates the importance of control variables related to ability and family characteristics - after these are controlled for, the music coefficient decreases substantially. Interestingly, adding control variables for individual characteristics and household income does not alter the music effect.

Future research should therefore consider the relevance of ability and parents when studying the impact of music.

Musical activity matters for educational outcomes. Given that music has the potential to increase human capital formation, equal access to music is an important economic issue. The Current Population Survey, conducted by the U.S. Census Bureau, has recently shown that only about 12 percent of U.S. adults play music. Music making seems all the more exclusive when this 12 percent of musically active is categorized by race, education, and household income. Indeed, the likelihood of playing music in adulthood almost doubles for white respondents, respondents with college degrees, and adults with higher incomes (NEA, 2008). In view of the strong reputation that music making continues to enjoy, these low rates of music participation are surprising.

Considering the low participation rates observed in the U.S. census data, a central policy conclusion relates to the public provision of this good. Despite less expensive instruments and better access to music lessons, participation rates remain relatively low. As musical activity plays such an important role in the determination of cognitive skills, equal access to musical instruments and lessons should be promoted. Initially sung or played on rudimentary instruments, music activity has become increasingly complex with the development of instruments, notation, and music theory (Hagel, 2010). The exclusiveness of music makes itself felt for music lessons, nowadays still a matter of financial resources and highly exclusive (Sadie, 1990). A potential means of providing every individual with the opportunity to learn to play an instrument is the promotion of obligatory music lessons in schools, in addition to ordinary music theory lessons that are subsidized by the state. Initial attempts to integrate such instrument lessons have been made in several U.S. and German schools. However, most programs have not featured both instrumental lessons and music theory lessons but one or the other. This paper stresses the importance of musical activity but does not discount the relevance of music listening and music theory. Future classrooms should include both instrument lessons and music theory.

References for Chapter 4

Angrist, J. D. and Pischke, J.-S. (2009), Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

Becker, G. S. (1993), Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education 3rd edition, Chicago, University of Chicago Press.

Bilhartz, T. D., Bruhn, R. A., and Olson, J. E. (2000), The Effect of Early Music Training on Child Cognitive Development, *Journal of Applied Developmental Psychology* 20(4), 615-636.

Boes, S. and Winkelmann, R. (2006), Ordered Response Models, Allgemeines Statistisches Archiv- Journal of the German Statistical Society 90, 167-181.

Borjas, G. J. (2000), Foreign-Born Teaching Assistants and the Academic Performance of Undergraduates, *American Economic Review* 90(2), 355-359.

Ceci, S. J. and Williams, M. (1997), Schooling, Intelligence, and Income, *American Psychologist* 52(10), 1051-1058.

Chabris, C. F. (1999), Prelude or requiem for the "Mozart effect", *Nature* 400(6747), 826-827.

Cooke, L. P. (2003), A comparison of initial and early life course earnings of the German secondary education and training system, *Economics of Education Review* 22(1), 79-88.

Costa-Giomi, E. (1999), The effects of three years of piano instruction in children's cognitive development, *Journal of Research in Music Education* 47, 198-202.

Darnell, A. and Darren, E. S. (1997), The impact of protestant fundamentalism on educational attainment, *American Sociological Review* 62(2), 306-315.

Dustmann, C., Puhani, P.A., and Schoenberg, U. (2012), The Long-term Effects of School Quality on Labor Market Outcomes and Educational Attainment, Centre for Research and Analysis of Migration Discussion Paper Series No 08/12, London.

Dustmann, C. (2004), Parental background, secondary school track choice and wages, *Oxford Economic Paper* 56, 209-230.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995), Increased cortical representation of the fingers of the left hand in string players, *Science* 13, 305-307.

Edelson, R. J., and Johnson, G. (2003), Music Makes Math Meaningful, *Childhood Education* 80(2), 65-70.

Forgeard et al. (2008), The Relation Between Music and Phonological Processing in Normal-Reading Children and Children with Dyslexia, *Music Perspective* 25(4), 383-390.

Frick, J. R. and Samberg, O. G. (2007), Biography and Life History Data in the German Socio Economic Panel, DIW Data Documentation 36.

Garlin, F. and Owen, K. (2006), Setting the tone with the tune: A meta-analytic review of the effects of background music in retail settings, *Journal of Business Research* 59(6), 755-764.

Gardiner, M. F., Fox, A. K, Nowles, F., and Jeffrey, D. (1996), Learning improved by arts training, *Nature* 381(6580), 284.

Gottfredson, L. S. (2002), Gottfredson's theory of circumscription, compromise, and selfcreation. In D. Brown (Ed.), Career choice and development (4th ed., pp. 85-148). San Francisco: Jossey-Bass.

Gottfredson, L. S., and Lapan, R. (1997), Assessing gender-based circumscription of occupational aspirations. *Journal of Career Assessment* 5, 419-441.

Hagel, S. (2010), Ancient Greek Music: A New Technical History, Cambridge University Press.

Hall, C. (2005), Gender and boys' singing in early childhood, *British Journal of Music Education* 22, 5-20.

Hetland, L. (2000), Learning to make music enhances spatial reasoning, *Journal of Aesthetic Education* 34, 179-238.

Hille, A. and Schupp, J. (2013), How learning a musical instrument affects the development of skills, IZA Discussion Papers 7655.

Hyde, K. L., Lerch J., Norton A., Forgeard M., Winner E., Evans, A. C., and Schlaug, G. (2009), Musical training shapes structural brain development, *Journal of Neuroscience* 29, 3019–3025.

Johnson, C. M. and Memmott, J. E. (2006), Examination of Relationships between Participation in School Music Programs of Differing Quality and Standardized Test Results, *Journal of Research in Music Education* 54(4), 293-307

Kaempfe, J, Sedlmeier,, P., and Renkewitz, F. (2010), The impact of background music on adults listeners: A meta-analysis, *Psychology of Music*, 39(4), 424-448.

Landels, J. G. (2002), Music in Ancient Greece and Rome, Routledge.

Lee, D. J., Chen, Y., and Schlaug, G. (2003), Corpus callosum: Musician and Gender effects, *Neuroreport* 14, 205-209.

Lechner, M. (2009), Long-run labour market and health effects of individual sports activities, *Journal of Health Economics* 28(4), 839-854.

Lehrer, E. L. (1999), Religion as a Determinant of Educational Attainment: An Economic Perspective, *Social Science Research* 28(4), 358-379.

Monaghan, P., Metcalfe, N. B., and Ruxton, G. D. (1998), Does practice shape the brain?, *Nature* 394, 434.

Morrison, S. J. (1994), Music students and academic growth, *Music Educators Journal* 80(2), 33-36.

Nantias, K. M. and Schellenberg, E. G. (1999), The Mozart effect: An artifact of preference, *Psychological Science* 10, 370–373.

NEA (2008), Survey of Public Participation in the Arts, National Endowment for the Arts, Research Report No. 49.

Neisser, U., Boodoo, G., Bouchard, Jr. T., Boykin, A. W. (1996), Intelligence: Knowns and unknowns, *American Psychologist* 51(2), 77-101.

Nantais, K. M. and Schellenberg, E. G. (1999), The Mozart Effect: An Artifact of Preference, *Psychological Science* 10(4), 370-373.

Pantev, C. , Oostenveld, R., Engelien, A. Ross, B. Roberts, L. E., and Hoke, M. (1998), Increased auditory cortical representation in musicians, *Nature* 392, 811-814.

Pelosi, F. (2010), Plato on Music, Soul and Body, Cambridge University Press.

Petrill, S. A., Lipton, P. A., Hewitt, J. K., Plomin, R., Cherny, S. S., Corley, R., and DeFries, J. C. (2004), *Developmental Psychology* 40(5), 805-812.

Pfeifer, C. and Cornelißen, T. (2010), The impact of participation in sports on educational attainment- New evidence from Germany, *Economics of Education Review* 29, 94-103.

Plomin, R., Fulker, D. W., Corley, R., and DeFries, J. C. (1997), Nature, Nurture, and Cognitive Development from 1 to 16 Years: A Parent-Offspring Adoption Study, *Psychological Science* 8(6), 442-447.

Pont, G. (2004), Philosophy and Science of Music in Ancient Greece: The Predecessors of Phythagoras and their Contribution, *Nexus Network Journal* 6(1), 17-29.

Rainbow, B. (1988), Music in educational though and practice: A survey from 800 BC, Boethius Press.

Rauscher, F. H., Haw, G. L., and Levine, L. J.(1997), Music training causes long-term enhancement off preschool children's spatial-temporal reasoning, *Neurological Research 19*, 2-8.

Rauscher, F. H., Shaw, G. L., and Ky, K. N. (1993), Music and spatial task performance, *Nature* 365, 611.

Sadie, J. A. (1990), Companion to Baroque Music, University of California Press.

Sandner, W. (1992), The effects of ethnicity and religion on educational attainment, *Economics of Education Review* 11(2), 119-135.

Sandner, W. and Krautman, A. C. (2007), Catholic schools, dropout rates and educational attainment, *Economic Inquiry* 33(2), 217-233.

Schellenberg, E. G. (2011), Examining the association between music lessons and intelligence, *British Journal of Psychology* 102(3), 283-302.

Schellenberg, E. G. (2006), Long-term positive associations between music lessons and IQ, *Journal of Educational Psychology* 98(2), 457-468.

Schellenberg, E. G. (2004), Music lessons enhance IQ, Psychological Science 15, 511-514.

Schellenberg, E. G. (2001), Music and nonmusical abilities, Annals of the New York *Academy of Sciences* 930, 355-371.

Schlaug, G., Jaencke, L. Huang, Y., Staiger, J. F., and Steinmetz, H. (1995), Increased corpus callosum size in musicians, *Neuropsychologica* 33(8), 1047-1055.

Schmidt, F. L. and Hunter, J. E. (1998), The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research finding, *Psychological Bulletin* 124(2), 262-274.

Schumacher, R. (2009). “Pauken mit Trompeten. Lassen sich Lernstrategien, Lernmotivation und soziale Kompetenzen durch Musikunterricht fördern?” Bildungsforschung Band 32, Bundesministerium für Bildung und Forschung (BMBF).

Sergeant, D. and Thatcher, G. (1974), Intelligence, Social Status and Musical Abilities, *Psychology of Music* 2(2), 32-57.

Sichivitsa, V. O. (2007), The influences of parents, teachers, peers and other factors on students’ motivation in music, *Research Studies in Music Education* 29(1), 55-68.

Southgate, D. E. and Roscigno, V. J. (2009), The Impact of Music on Childhood and Adolescent Achievement, *Social Science Quarterly* 90(1), 4-21.

Vaughn, K. (2000), Music and Mathematics: Modest Support for the Oft-Claimed Relationship, *Journal of Aesthetic Education* 34(4), 149-166.

Wechsler, D. (1991), WISC-III: Wechsler intelligence scale for children, third edition, San Antonio, TX: The Psychological Corporation.

Winner, E., T. G. Goldstein, and Vincent-Lancrin, S. (2013), Arts for Arts’s Sake? The Impact of Arts Education, OECD Publishing.

Wooldidge, J. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press.

Zdzinski, S. F. (1996), Parental involvement, selected students attributes, and learning outcomes in instrumental music, *Journal of Music Education* 44(1), 34-48.

Appendix A4.1: Variables and Coding

Table 4.7. Variables and Coding

Variable	Description
Music variables	
Music active	Musically active (dummy)
Music paid	Received paid music lessons (dummy)
Music often	Daily musically active (dummy)
Music seldom	Weekly/Monthly musically active (dummy)
Music early childhood	Started Music between age 0-6
Music childhood	Started Music between age 7-12
Music youth	Started Music between age 12-17
Music alone	Plays music alone (dummy)
Music with	Plays music with others (dummy)
School variables	
School track	Attends “No” (0) “Low” (1) “Intermediate” (2) “High” (3)
Track no	Middle school dropouts (dummy)
Track low	Low middle school track (dummy)
Track med	Med middle school track (dummy)
Track high	High middle school track (dummy)
Track rec	Recommended school track “Low” (1) “Intermediate” (2) “High” (3)
Personal characteristics	
Male	Male student (dummy)
Sibling	Existence of siblings (dummy)
Immigration	Migration Background (dummy)
Town size	“Very small” (0) “Small” (1) “Large” (2) “Very Large” (3)
Move child	Moved from cities during childhood (dummy)
Parental variables	
Mother higher educated	Mother completed higher education(dummy)
Mother med. Educated	Mother completed intermediate education(dummy)
Father higher educated	Father completed higher education(dummy)
Father med. Educated	Father completed intermediate education(dummy)
Loghhincome 2000	Log Household income in 2000
Parents push	No (1) Some (2) Strong (3) Very strong (3)
Parents help	Parents help (dummy)
Parents problems school	Problems with parents because school (dummy)
Parents school nights	Parents attend school nights (dummy)
Parents teachers	Parents consult teachers (dummy)
Instrumental variables	
Church affiliation	Goes to church (dummy)
Protestant	Member of the Protestant church (dummy)

Appendix A4.2: Data Appendix

Uniquely, BIOAGE17 features a rich set of retrospective, current and prospective questions related to educational path, leisure activities, and household characteristics; including information on past track recommendation, last school track, music activity, type of music lessons, intensity of music activity, begin of music activity, mode of music activity, gender, the amount of siblings, town size, county of residence, and an indicator for mobility during childhood. Furthermore, BIOAGE17 contains a number of statements with respect to the current relationship with parents and parental involvement in school- and leisure-activities.

When reducing the sample to adolescents without any missing information on music or school track at age 17, the final sample contains information on 3,222 adolescents at age 17. Few youths have missing observations with respect to the school track at age 17 or the confounding variables on individual or family level- including education of parents and migration status- that even if the sample was to be reduced to observations without any missing information of controls on individual and family level, the dataset would still comprise 3,183 adolescents; less than 5 percent reduction from the initial BIOAGE17 sample. Unfortunately, the information on track recommendation does not exist for around 20 percent of the sample. For this reason, the latest track attended at age 17 is also considered.

BIOAGE17 contains no information on the highest education of parents, migration background and household income. Information on the education of parents is important to proxy the cognitive level of parents; its inclusion also serves to eliminate potential ability bias if only children of higher ability or intellectual household engage in music. Data on migration background can be used to control for ethnicity-related access to music and race-specific school paths. Reports on household income are relevant to exclude income-related effects on music activity and school tracks. High income households do not face financial constraints with respect to costly music instruments and music lessons. At the same time families with more resources are able to invest more into children's education, increasing a child's probability of completing higher education.

Fortunately, this information was available in other SOEP datasets. The Biography Information for the Parents of SOEP-Respondents (BIOPAREN) provides information on the highest education of both parents and the Generated and Status Variables from SOEP for Foreigners and Migrants (BIOIMMIG) contains answers regarding past and current immigration status. After adding that information to the BIOAGE17 dataset, the sample consisted of 3,265 adolescents. However, 43 of these respondents did not respond to either music-related question or school track at age 17.

Information on household income was derived from the main SOEP dataset that collects yearly information on corrected monthly net household income. Unfortunately, when combining the net household income before 2000 - when the adolescent were younger than 17 years- with the BIOAGE17 dataset, more than 50 percent of the sample contains missing information on income. The amount of missing data on net household income is the smallest if only the data on net household income in 2000 is combined with BIOAGE17. Then, the initial sample is reduced by less than 30 percent. Still, due to the sizable reduction of observations with net income, the empirical analysis with information on household resources is only used for robustness checks.

For educational outcomes, this paper reports track recommendations and last track at age 17. Unlike track recommendations, this record not only includes traditional middle schools but also alternative types of middle school. For instance, some 17-year-olds last attended German comprehensive schools – *Gesamtschule* – that include all types of students and are comparable to US high schools. Others, by contrast, last attended specialized secondary schools - *Fachoberschule*. Although it is very likely that the 17-year-old students surveyed proceeded to complete higher middle school, this assumption was confirmed using statements on the date of middle school completion and, preferably, a middle school diploma. In this study, school track at age 17 is categorized as follows: 0 for “dropout”, 1 for “low”, 2 for “intermediate”, and 3 for “high” middle school. In addition, this paper uses binary variables for higher middle school (“track high”), intermediate middle school (“track med”), and low middle school (“track low”). If no track is completed or attended, the respondent is marked as a dropout (“track no”).

The instruments used for the robustness checks are church affiliation and membership in the Protestant church. BIOAGE17 includes reports on the regularity of church visits through which church affiliation is proxied. Church affiliation is a dummy variable that is one, if the respondents go to church on at least seldom basis. Protestant is a dummy variable that is one if the respondent is a member of the Protestant church. Unfortunately, this question included a substantial amount of missing observation, around 60 percent. Therefore, the results are to be interpreted cautiously.

Appendix A4.3: The German middle school system tracking system

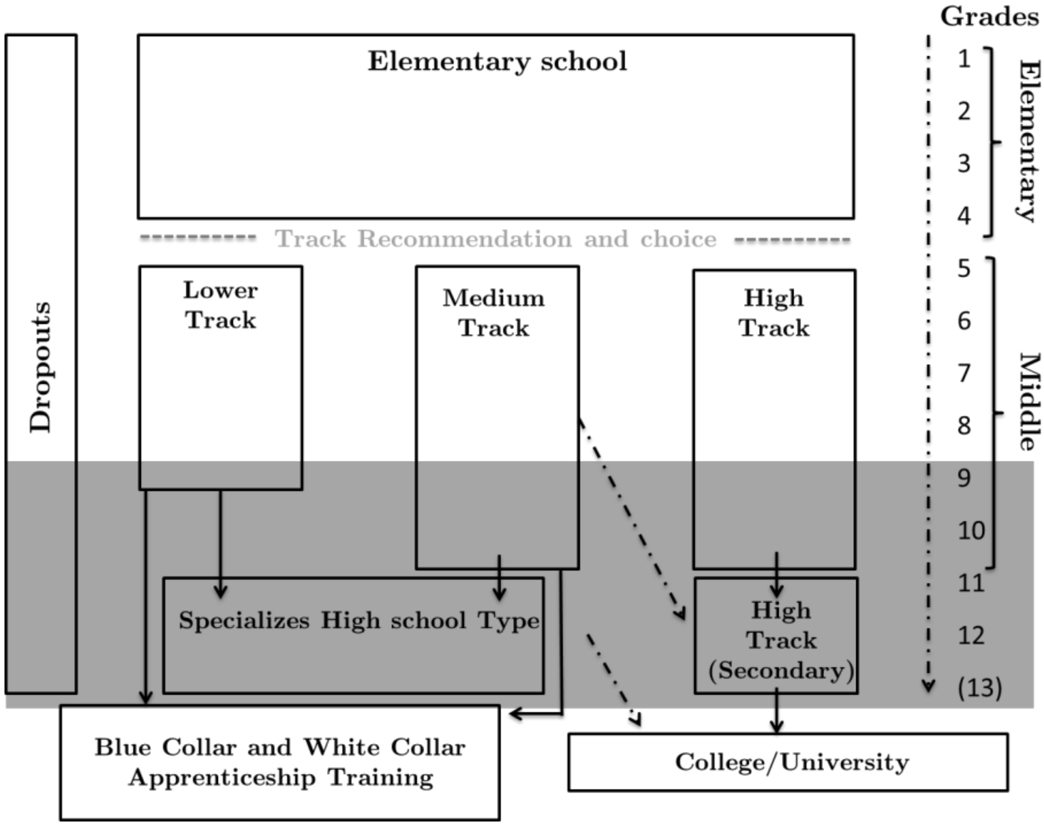


Figure 4.2. Overview of the German middle school tracking system.

Note: This figure, based on Dustmann et al. (2012), provides detailed information on the German tracking system. After children have completed four years of elementary school, teachers make track recommendations for each child. Depending on state law, these recommendations are either binding or not binding. In some cases, a track recommendation can only be overridden through additional examination. Thereafter, students are assigned to middle schools on either the lower track, the medium track, or the higher track. Children in the lower or intermediate tracks are likely to enter blue or white collar apprenticeship programs, whereas children in higher tracks are likely to enrol in the higher secondary track and, thereafter, college or university. However, the German school system offers several opportunities to switch educational paths during and after middle school, denoted by the dashed arrows. The BIOAGE17 sample is in the grey area, where adolescents are either at the end of tenth grade in middle school, in specialized high school, in apprenticeship training, or have dropped out. Some may already be employed and have completed their education. For these individuals, this study only observes the last school track attended.

After grade four of elementary school, children around age ten receive track recommendations from teachers; lower-, intermediate- or higher track recommendations. These recommendations are either binding or non-binding in dependence of German state laws and residence. According to Dustmann et al. (2012), the associated middle school types differ with respect to teaching content, peers, teacher quality, and teaching intensity. The curriculum of the traditional three types of schools distinctively prepares adolescents for different educational paths and careers. The highest middle school type – *Gymnasium* - aims to prepare students for higher academic institutions for nine years. Intermediate middle schools - *Realschule* - and lower middle schools – *Hauptschule* - are only five to six years long and prepare student for apprenticeships in blue- and white-collar occupations. Track recommendations by teachers are either recommendations for Gymnasium, Realschule or Hauptschule. This study labels these recommendations ("track rec") along the school types; where 1 equals a "low", 2 equals a "intermediate" and 3 equals a "high" recommendation. Track recommendations, however, are not only interesting to predict the educational paths of children in the future. Since teachers recommend students to different school tracks after four years of elementary schools, these recommendations are also interpretable as a third party evaluation of a child's cognitive skill

Appendix A4.4: German Federal states and binding and non-binding recommendations

To relate music to long-run outcomes, the main outcome variable used is school track at age 17. The comparison of track recommendation and school track at age 17 is useful if the track recommendations are non-binding. In this case, parents can enroll their children in a higher or lower track than recommended, and thus the school track at age 17 may very well differ from the initial recommendation if students complete a type of school other than the one suggested. Even if track recommendations are binding, it is very likely that some children with low track recommendations will complete higher school tracks, and vice versa. The German tracking system incorporates two main sources of “second chances” (Dustman et al., 2012). On the one hand, it is possible to horizontally shift between middle school types during middle school. On the other hand, following completion of one school type, students can vertically shift to higher educational tracks. It is through these second chances that adolescents, even if the school recommendations are binding, may attend a different middle school from the one initially recommended.

Table 4.8 informs about the regulations of track recommendations for each German federal state. Additionally, this table informs how the track recommendation is enforced. In case of non-binding recommendations, parents are able to send their children to any kind of school. If the recommendations are binding, students have to go through additional exams or have a certain trial period and exams if they attend a higher school than recommended.

Table 4.8. Regulations of track recommendations by state

Federal State	**Binding**	**Enforcement**
Baden-Württemberg	Yes	Test
Bayern	Yes	Trial and test
Brandenburg	Yes	Schools decision and test
Bremen	Yes	Schools decision, Interview
Saarland	Yes	Test
Sachsen	Yes	Test
Sachsen-Anhalt	Yes	Test
Schleswig-Holstein	Yes	Test
Thüringen	Yes	Test
Hamburg	No	
Hessen	No	
Mecklenburg-Vorpommern	No	
Niedersachsen	No	
Nordrhein-Westfalen	No	
Rheinland-Pfalz	No	
Berlin	No	

Table 4.8 shows that about halve of the German federal states enforce binding track recommendations. If track recommendations are binding, students, in most of the cases, have to pass additional tests if they want to enter a higher middle school track than recommended.

Appendix A4.5: Track recommendation and track at age 17

Table 4.9 provides information on track recommendation and school track at age 17 for musicians and non-musicians in Panel A. Furthermore, it reports cross tabulations of track recommendation and school track at age 17 for musicians and non-musicians in Panel B

Table 4.9. School track at age 17, track recommendations and music

Panel A: Shares of non-musicians and musicians in school tracks and track recommendations								
	Non-musicians				Musicians			
	Drop	Low	Med.	High	Drop	Low	Med.	High
School track at age 17	0.008	0.150	0.441	0.401	0.003	0.044	0.251	0.702
Track recommendation	-	0.221	0.382	0.397	-	0.084	0.240	0.676

Panel B: Cross tabulation of track recommendation and school track at age 17								
	Non-musicians				Musicians			
	Drop	Low	Med.	High	Drop	Low	Med.	High
Track rec. Low	0.007	0.410	0.490	0.093	0.000	0.286	0.518	0.196
Track rec. Med	0.001	0.073	0.647	0.279	0.000	0.038	0.515	0.447
Track rec. High	0.001	0.005	0.177	0.817	0.000	0.004	0.105	0.891
p-value	0.092	0.000	0.000	0.000	.	0.000	0.000	0.000

Note: Panel A reports the shares of non-musicians and musicians in a given track at age 17 (dropouts, lower track, intermediate track, or higher track) and the shares of initial recommendations after elementary school (low, intermediate or higher track recommendation). The sample consists of 3,222 observations (2,470 non-musicians and 752 musicians) of last tracks attended. Reports on track recommendations are only available for 2,595 respondents (1,931 non-musicians and 664 musicians). Panel B displays cross tabulations of track recommendation and school track attended at age 17 for 2,595 observations (1,931 non-musicians and 664 musicians). The last row in Panel B reports the p-value for the hypothesis of no significant relationship between track recommendation and school track at age 17.

Notably, Panel A in Table 4.9 reports a significantly larger share of musicians in higher middle school tracks relative to the musically inactive; 70 percent of musicians and only 40 percent of non-musicians attended higher middle school tracks. A larger share of non-musicians than musicians last attended intermediate school or were in lower middle school tracks; some 25 percent of the musicians and 44 percent of the non-musicians were in intermediate-school tracks, and four percent of the musicians and 15 percent of the non-musicians were in lower-middle school tracks. The number of dropouts is very small for both groups. Many more musicians than non-musicians receive higher track recommendations, approximately 68 percent of all musicians compared to 40 percent of non-musicians. Musically inactive respondents are more likely to receive lower or intermediate recommendations.

The cross tabulation of school track at age 17 and track recommendation in Panel B provides information on this relationship for the groups of non-musicians and musicians. It is noteworthy that, although a larger share of musicians than non-musicians have high track recommendations, more musicians than non-musicians with lower or intermediate track recommendations manage to shift and remain in higher tracks. Approximately 20 percent of adolescents with initial low track recommendations attend higher middle schools tracks at age 17. Notably, 45 percent of musicians with intermediate school recommendations manage to shift to higher school tracks. Although these trends are also observable for the group of non-musicians, they are considerably smaller. In sum, musicians receive higher track recommendations and are more likely than non-musicians to make up for worse track recommendations by shifting to and remaining in higher middle school tracks by age 17.

Appendix A4.6: Specific recommendation using LPM

To examine the music-related transitions between specific recommendations, this paper provides regression results for linear probability models (LPM) in Table 4.10. Here, the regression coefficients indicate music-related effects on the likelihood of receiving a higher (Panel A) or at least an intermediate recommendation (Panel B). Given the relevance of individual and family variables observed in previous results, these variables are held constant in these regressions.

Table 4.10. Specific track recommendation and music

Music	1 **Active**	2 **Paid**	3 **Intensity**		4 **Begin**		5 **Peers**	
			Often	**Seldom**	**0-5 years**	**6-10 years**	**Alone**	**With**
<u>High track versus others (including individual and parental controls)</u>								
Coeff.	0.185***	0.205***	0.238***	0.168***	0.245***	0.243***	0.155***	0.207***
(s.e.)	(0.021)	(0.024)	(0.029)	(0.027)	(0.036)	(0.028)	(0.030)	(0.026)
N	2,567	2,567	2,567	2,567	2,567	2,567	2,567	2,567
R-squared	0.199	0.199	0.203	0.203	0.206	0.206	0.200	0.200
<u>High or mediocre versus low track (including individual and parental controls)</u>								
Coeff.	0.080***	0.072***	0.101***	0.076***	0.079***	0.081***	0.053***	0.098***
(s.e.)	(0.014)	(0.015)	(0.018)	(0.017)	(0.020)	(0.018)	(0.020)	(0.016)
N	2,567	2,567	2,567	2,567	2,567	2,567	2,567	2,567
R-squared	0.129	0.126	0.130	0.130	0.127	0.127	0.129	0.129

Note: This table reports various LPM estimates between the probability of a higher track recommendation and at least mediocre/high track recommendation and music indicators. Each of the five regressions is estimated using confounding variables at the individual level and family characteristics (including highest educational attainment of parents). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel A shows a positive and significant relationship between music and the likelihood of receiving a high track recommendation, with musical involvement increasing the probability of receiving a higher track evaluation by approximately 20 percentage points. Generally, the interpretations of the regression results for each music indicator (columns 1-5) are very similar to those using the threefold track variable. The effect of paid music lessons is slightly larger than that of the general music indicator. Higher track recommendations are more likely when students practice often rather than seldom. Music estimates are strongest when only children who began early are considered.

Panel B of Table A2 suggests a positive effect of music on the probability of at least receiving an intermediate recommendation. However, all coefficients are nearly halved relative to those in Panel A. According to Table 4, music is more helpful in the transition to receiving a higher track recommendation than in the transition to receiving at least an intermediate recommendation.

To distinguish the effects of music on the transitions between particular tracks, this paper reports the estimations of linear probability models of the impact of music on specific track forms separately. Table A3 provides correlations between the different music indicators and the probability of attending a higher track (Panel A), attending at least an intermediate track (Panel B) and attending at least one additional track beyond elementary school (Panel C). These results provide insights into music-related transitions between each track. In each model, control variables for personal and parental characteristics are included.

Table A4.11. Specific school track at age 17 and music

Music	1 **Active**	2 **Paid**	3 **Intensity**		4 **Begin**			5 **Peers**	
			Often	**Seldom**	**0-5 years**	**6-10 years**	**11-17 years**	**Alone**	**With**
(i) High track versus others (including individual and parental controls)									
Coeff.	0.196***	0.221***	0.239***	0.176***	0.284***	0.214***	0.119***	0.201***	0.192***
(s.e.)	(0.019)	(0.021)	(0.025)	(0.024)	(0.029)	(0.026)	(0.030)	(0.026)	(0.023)
N	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183
R-squared	0.229	0.229	0.231	0.231	0.233	0.233	0.233	0.229	0.229
(ii) High /mediocre versus low track/ dropout (including individual and parental controls)									
Coeff.	0.066***	0.068***	0.068***	0.066***	0.085***	0.073***	0.049**	0.044***	0.080***
(s.e.)	(0.010)	(0.010)	(0.015)	(0.012)	(0.011)	(0.012)	(0.019)	(0.015)	(0.012)
N	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183
R-squared	0.123	0.122	0.123	0.123	0.123	0.123	0.123	0.124	0.124
(iii) High /mediocre/low versus dropout (including individual and parental controls)									
Coeff.	0.002	0.005***	0.005***	0.003	0.004***	0.002	0.002	0.001	0.003
(s.e.)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.004)	(0.004)	(0.004)	(0.003)
N	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183	3,183
R-squared	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007

Note: This table reports various OLS estimates of the probabilities of having attended (i) a higher track, (ii) at least a mediocre track, (iii) of not dropping out, given music indicators. Each of the five regressions was estimated using confounding variables at the individual level and family characteristics (including highest educational attainment of parents). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In Table A3, Panel A reports a significant and positive effect of music on the probability of attending a higher middle school track by age 17. Specifically, musical activity increases the likelihood of being in a higher track by approximately 20 percentage points, irrespective of the music model considered and of inclusion of personal and family control variables. Overall, the results for the different music models are very similar to those reported above. The effect of music is slightly stronger when music lessons are paid for. Musical activity has a stronger effect when it is frequent than when it is seldom. The earlier that musical instruction begins, the larger the effect. Only peers appear not to matter in the transition to higher middle school tracks.

Panel B shows the regression results for the probability of attending at least intermediate middle school. In this specification, all coefficients are approximately one-third the size of those in Panel A, but all are significant. In Panel C, which relates the probability of attending at least one further track after elementary school to musical activity, the coefficients are either not significant or marginal.

Appendix A4.7: Further robustness checks

This paper further estimate generalized ordered probit (GOP) models, following Boes and Winkelmann (2006), to test the validity of the previous results and show that they are not biased because of miscalculations associated with functional form. GOP models were selected over ordinary ordered probit models because the former relax the restrictions that the marginal probability effects must be constant and may only change sign once from the largest to the smallest outcome. The marginal effects in GOP models do not treat these threshold values as constant but make them dependent on the regressors, which allows for greater flexibility in the estimation of the marginal probability effects (Pfeifer and Cornelißen, 2010). A further advantage of estimating GOP models is improved interpretability of the results. Whereas OLS results only denote changes in standard deviations, GOP results indicate predicted marginal probabilities of being in a certain outcome group due to the music variable.

The primary purpose of the GOP estimation is to indicate whether the previous OLS or Ordered Probit results suffer from bias related to the linearity assumption. Another advantage of the GOP estimation concerns the interpretation of the results. Whereas OLS yields differences only in terms of standard deviations, the GOP estimates provide probabilities of obtaining various levels of educational attainment. Table 4.12 reports predicted probabilities of the means of track recommendation and school track at age 17, as explained by musical activity and the control variables at the individual, parental, income, and state levels.

Table 4.12. Generalized Ordered Probit

	Track recommendation			School track at age 17			
	Low	**Intermediate**	**High**	**No**	**Low**	**Intermediate**	**High**
Prob. at means	18.6%	34,5%	46.9%	0.6%	12.5%	39.7%	47.1%
Music activity	-9.1%***	-12.0%***	**21.1%*** **	-0.4%***	-2.0%***	-16.8%***	**23.9%*** **
	(0.010)	(0.023)	(0.024)	(0.001)	(0.008)	(0.021)	(0.022)
Individual controls		Yes				Yes	
Parental controls		Yes				Yes	
Income controls		Yes				Yes	
State fixed-effects		Yes				Yes	
N		2101				2648	
Log likelihood		-1885.85				-2321.08	

Note: Probabilities and effects computed for means; z-values are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The estimates of the GOP model support the previous OLS results under the CIA. Music has a significant and positive effect on educational attainment. Musical activity increases the predicted probability of receiving a higher track recommendation after elementary school by 21 percent, while lower and intermediate recommendation probabilities decrease significantly in the presence of musical activity. Musical activity increases the probability of being in a higher school track at age 17 by 24 percent and significantly reduces the probability of being in an intermediate school track.

Propensity score matching (PSM) has gained increasing interest as an empirical tool. Considered similar to saturated regressions, propensity score matching is used to check the results of the regressions and whether the CIA holds. In particular, this paper applies nearest neighbor matching (NMM) and kernel-based matching (KM). Both methods are used, as they are likely to differ in terms of bias and efficiency; the NMM selects the closest non-participant to each participant to construct a pair of treated and non-treated individuals, and the non-parametric KM includes all non-participants and uses weights to reduce the influence of more distant propensities (Caliendo and Kopeinig, 2005).

To validate the results of OLS under the CIA, this study estimates the effect of musical activity on educational attainment, using PSM. As the results from PSM are unlikely to be significantly different from a nearly saturated model, the results from PSM indicate whether all relevant information was held constant in the previous OLS model. This study considers two types of matching methods, NNM and KM, as robustness checks. Both matching methods are employed because they differ significantly with respect to bias and efficiency. Table 4.13 reports the estimates of both PSM methods for both track recommendations and track attended at age 17.

The results of both matching methods in Table 4.13 suggest no significant differences from the OLS model under the CIA. Interestingly, the estimated average treatment effect of the treated (ATT) in the PSM model is nearly identical in size and significance, irrespective of which matching method used, as the previously observed music effect using OLS. This suggests that the OLS model under the CIA is effectively saturated. According to the results of the PSM strategy, musical activity increases track recommendation and school track at age 17 by approximately 0.2.

Table 4.13. Robustness checks using Propensity Score Matching

	Track recommendation		School track at age 17	
Model	PSM	PSM	PSM	PSM
Method	Nearest Neighbor matching	Kernel based matching	Nearest Neighbor matching	Kernel based matching
ATT (music acitivity)	0.279***	0.298***	0.292***	0.289***
(s.e.)	(0.047)	(0.045)	(0.055)	(0.027)
Individual controls	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes
Parental controls	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Mean absolute st. difference	30.792	3.840	31.026	3.361
Rosenbaum test	Yes	Yes	Yes	Yes
N	1445	3222	1663	3222

Note: ATT effects of PSM were estimated on basis of a probit model. The region of common support is [0.054, 0.639] and the final number of block equals 6. Balancing property is satisfied. Bootstrapped standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.