The London School of Economics and Political Science

Performance Pay in Academia: Effort, Selection and Assortative Matching

Erina Ytsma

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, August 2015

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 38,930 words.

Acknowledgements

I would not have been able to write this thesis without the love, support, patience and understanding of Baran, Jesus and the rest of the family. I would also like to thank my parents, brother and grandmothers for their continued interest in what I do and where I am, their love and support. And I would like to thank my friends for knowing exactly when I do and when I do not want to talk PhD.

I am grateful for having been given the opportunity to do a PhD at LSE, for the financial support I have received from the ESRC and LSE to be able to do so, and for the time and energy so many professors and fellow PhD's have selflessly given to me during my PhD. A special mention also for the admin and support staff, who have always been so kind and helpful.

I want to express my sincere gratitude to De Gruyter publishers for graciously providing me with their Deutscher Gelehrten-Kalender database, and gratefully acknowledge all the help and support I have received in this process from Axel Schniederjuergen.

And last, but anything but least, I am grateful to Oriana Bandiera, Tim Besley and Andrea Prat for their supervision and advice, without which I would not have been able to write this thesis.

Abstract

This thesis studies the effect of performance pay on effort, selection and matching assortativeness in academia, using the introduction of performance pay in German academia as a natural experiment and employing a newly constructed data set encompassing the affiliations and productivity of the universe of academics in the country.

I estimate the pure effort effect in a difference-in-differences framework comparing the productivity of cohorts that started their first tenured position just before the reform, and consequently do not receive performance pay, with those starting their first tenured position after the reform, and therefore do receive performance pay. I find that the effort effect is economically large; amounting to a 35% increase in academic productivity relative to the pre-reform productivity in the control group.

I estimate the selection effect by analysing the rate at which academics of different productivity levels switch to the performance pay scheme and by exploiting the fact that the old and new wage scheme compare differently for academics at different ages, which gives rise to selection incentives that are inversely related to age. I find that more productive academics are more likely to select into performance pay, and that this effect is stronger for younger academics.

The empirical framework to study matching assortativeness is informed by a simple matching model in which I show that performance pay increases positive assortative matching if there are positive productivity spillovers, and that this increase is larger if complementarities are stronger. I test this hypothesis in a difference-in-differences framework using a measure of complementarity strength as a continuous treatment variable and find that assortative matching increases more in fields with stronger complementarities, thus providing empirical evidence that performance related pay increases positive assortative matching. This effect is large; amounting to a two- to threefold increase in positive assortative matching.

Contents

List of Figures

List of Tables

1 Introduction

Academia constitutes an important economic sector, representing both a large direct economic value and sizeable exports. In the UK, for instance, the sector amounts to an estimated direct value of 59 billion Euros and accounts for around 4.5 billion Euros in exports (McCormack et al. 2014). Furthermore, universities are important producers of research, particularly basic research, accounting for about 50% of basic research in the United States in the late 90's (Lach and Schankerman 2008 citing NationalScienceBoard 2000). Academic research, in turn, has been shown to have significant local spillover effects (Kantor and Whalley 2014, Agrawal 2001), as well as increase total factor productivity growth (Adams 1990). It is therefore important to understand what determines the performance of universities.

Human capital is the primary input factor in academia, so human resource management should have first order effects in the sector. Studying the effects of human resource management on the performance of universities is therefore both relevant for the debate about university governance, and for our understanding of the effects of human resource management systems in the workplace at large. Accordingly, a number of studies have analysed the effect of various aspects of human resource management on university productivity, from practices in general (McCormack et al. 2014), to more specific aspects such as incentives in the form of inventor royalty shares (Lach and Schankerman 2008, 2004). Aghion et al. (2010) take a step back, to factors influencing university governance, and study how the level of autonomy that a university enjoys and the amount of competition it faces affects its productivity.

This thesis focuses on one particular aspect of human resource management; pay, specifically performance pay. Performance pay is both widespread and becoming ever more prevalent. Lemieux et al. (2009) for instance show, using PSID data, that the incidence of performance pay for US salaried workers increased from less than 45% in the '70's to more than 60% in the '90's. Incentive systems have been widely studied (see e.g. Lazear and Oyer 2012, Bloom and Reenen 2011, Oyer and Schaefer 2011, Lazear and Shaw 2007, for excellent reviews of the literature), and many studies have reported positive effects of performance pay on productivity through increased effort (cf. Hossain and List 2012, Muralidharan and Sundararaman 2011, Boly 2011, Dohmen and Falk 2011, Carpenter et al. 2010, Bellemare et al. 2010, Lavy 2009, Bandiera et al. 2005, Shearer 2004, Lazear 2000a). The first question this thesis aims to answer is whether performance pay has a positive effort effect in academia as well.

Because academics might be particularly intrinsically motivated (McCormack et al. 2014), and because this intrinsic motivation might be crowded out by extrinsic incentives (Dickinson and Villeval 2008, Besley and Ghatak 2005), it is not ex ante obvious that performance pay increases effort in academics. The effect of performance pay on teachers' effort has, for instance, both been found to be positive and significant (Muralidharan and Sundararaman 2011) and insignificant (Glewwe et al. 2010). Lach and Schankerman (2008) and (2004) find that higher inventor royalty shares are associated with larger licensing revenues using university level data. The first part of this thesis adds to our understanding of incentives in universities by showing that performance pay increases the effort of academics in basic research, using individual level data.

As a next step, I study what kind of academics are attracted to performance pay. Lazear (2000a) finds that windshield installers who are more productive prefer piece-rate pay over a flat wage and consequently self-select into the former scheme and Dohmen and Falk (2011) present evidence of a similar selection effect in the lab. Leuven et al. (2011) show that higher ability students are more likely to select into tournaments in which they can win a larger prize. I find that in academia too, more productive academics are more likely to self-select into performance related pay. This is reassuring, since if lower productivity academics would be more likely to self-select into performance pay, the overall effect of performance pay on productivity in academia might be negative, and higher-powered incentives would not be a feasible means to increase academic output.

Finally, I study the effect of performance pay on the organisation of academics. Kremer (1993) shows that if there are complementarities in worker skill, workers match positive assortatively by skill. In turn, this causes output and wages to increase sharply with skill, and incomes to differ vastly across countries. Furthermore, we know from Legros and Newman (2002) that if complementarities are such that the output function is supermodular in skill, positive assortativeness increases total output. Since performance pay increases the utility that a worker derives from his output, workers have a (weakly) greater incentive to match positive assortatively by skill if there are complementarities in worker skill. If this holds true for academia as well, performance pay should increase positive assortative matching in academia too. This would increase total academic output if the academic output function is supermodular. A more positively assortative matching of academics by skill however also implies a less equal distribution of highly productive academics across universities. Because academic research has been shown to have significant local spillover effects through university-industry partnerships (Kantor and Whalley 2014, Agrawal 2001), a more positive assortative matching of academics by skill and consequently a greater concentration of academic research in a few places, would mean a concentration of such local spillovers in a limited number of areas. Furthermore, to the extent that academics who are productive researchers are also gifted educators, a more positive assortative matching would imply a less equal distribution of high quality tertiary education as well.

I use the introduction of a new performance pay scheme in German academia as a natural experiment to study the effect of performance pay on effort, selection and matching assortativeness in academia. This performance pay scheme was announced in 2002 and implemented in 2005. Any professorial contract signed or renegotiated after the pay reform necessarily falls under the performance pay scheme (Detmer and Preissler 2004). Before the reform, academics were paid according to an age-related pay scheme in which pay would increase every two years until the age of 49 (Hochschullehrerbund 2009). Under the age-related pay scheme, pay is thus effectively flat. The performance pay scheme involves a basic wage, which is lower than the wage under the age-related pay scheme for most ages (Hochschullehrerbund 2009, Oeffentlicher-Dienst 2004). On top of this basic wage however, professors can earn performance bonuses. Universities can award these bonuses to attract or retain a professor, for on-the-job performance, and for taking on management roles or tasks (BMBF 2002). Attraction or retention bonuses are generally awarded on the basis of qualifications and past performance in research and education and can be either temporary (at first) or permanent (Detmer and Preissler 2004, 2005). At most universities, the bonuses for on-the-job performance are distributed through, what are effectively, promotion tournaments, with promotion to a next hierarchy level associated with a pay rise (Lünstroth 2011, Kräkel 2006, Harbring et al. 2004). These bonuses are awarded on the basis of performance in research and education as well and they too can be permanent or temporary (Handel 2005). The on-the-job performance bonuses can range from 90 Euros to 2500 Euros[1] (Lünstroth 2011), while attraction or retention bonuses can be even larger, since most state laws do not stipulate an upper bound (Handel 2005). Accordingly, attraction bonuses account for the bulk of variable pay awarded under the performance pay scheme, while the management bonuses account for the smallest fraction (BMI 2007).

I constructed a new panel data set encompassing the personal details and information regarding the academic affiliations and research productivity of the universe of academics in Germany for the purposes of this research. The data set spans 15 years, from 1999 to 2013, and contains data on more than 55000 academics who, at some point in the 15 year timespan, held a tenured professorial position at a public university in Germany. It is these academics that I focus on in the empirical analyses in this thesis, because only tenured professors can earn bonuses under the performance pay scheme. I restrict attention to public universities, because I focus on the research dimension of academic output, and research output of higher education institutions other than universities is much smaller than that of universities.

In order to study the effort effect, I use the fact that any professorial contract renegotiated or signed after the implementation of the reform in 2005 falls under the performance pay scheme. Academics who first made tenure just before 2005 are thus paid according to the age-related pay scheme, while those who made tenure just after the reform fall under the performance pay scheme. If the timing of the tenure decision is exogenous we can then identify the effort effect of performance pay by comparing the research output of the academics who made tenure just after the reform with the productivity of academics who made tenure just before. I therefore estimate the effort effect in a difference-in-differences framework, using the academics who made tenure right after the reform as treatment group, and the academics who made tenure just before as control group. I find that performance pay increases productivity by 35%, about two-thirds of which can be ascribed to the tournament component of the performance pay scheme and a third to the piece-rate component of performance pay.

[1]In addition to monthly salary

As a next step, I present the results of a number of tests of the identifying assumptions. I find no evidence of pre-existing trends, which lends support to the identifying parallel-trends assumption of the difference-in-differences framework. Furthermore, I find no evidence undermining the assumption that the timing of the tenure decision is exogenous. A placebo difference-in-differences for instance shows that there is no differential change in productivity from before to after the reform between two cohorts of academics who made tenure before the reform. If higher or lower productivity academics were able to speed up their tenure clock, and if their productivity would change differentially once they have tenured, this would show up as a positive or negative effect of performance pay on the productivity of the cohort making tenure just before the reform (the placebo treatment group).

Apart from academics who make tenure after the reform and therefore necessarily fall under the performance pay scheme, academics can also self-select into the performance pay scheme. Academics who hold a tenured affiliation before the reform, and who are consequently paid according to the age-related pay scheme, switch to the performance pay scheme when they sign a new contract or renegotiate their existing one after the reform. Because pay increases with age in the age-related pay scheme, I would expect more productive academics to be more likely to switch to the performance pay scheme. I perform hazard rate and survival function analyses to test this hypothesis and find that more productive academics are more likely to select into the performance pay scheme indeed. Moreover, this effect is stronger for younger professors. Since pay increases with age in the age-related pay system, whereas the performance pay scheme pays a fixed basic wage, the performance pay scheme is relatively less attractive for older academics. Unless they are sufficiently productive so that their performance bonuses in the performance pay scheme will (more than) make up for the difference in basic wages, older professors will not self-select into the performance pay scheme.

Finally, I estimate the effect of performance pay on matching assortativeness. If there are complementarities in academic research, such that research output increases both in own productivity and in the productivity of colleagues, and if academics derive utility from their research output even in the age-related pay system, matching should be positively assortative by productivity. That is, we would expect to see highly productive academics work together in the same department, and similarly for low productivity academics. However, if academics also have idiosyncratic preferences, such as personal preferences and preferences regarding geographic location, matching would not be perfectly positive assortative. Some highly productive academics would then be affiliated with departments where the average productivity of the faculty is lower, and vice versa. In this case, performance pay should increase positive assortative matching by increasing the extent to which their utility depends on their output, provided there are spillovers so that output increases in the productivity of colleagues. Moreover, the increase in positive assortative matching should be larger if complementarities are stronger. It is this latter hypothesis that I use to identify the effect of performance pay on matching assortativeness and test in a difference-in-differences framework using a measure of the strength of complementarities as a continuous treatment variable.

A change in matching assortativeness implies a different distribution of academics across departments and hence a change in the composition of departments. I therefore study the two channels through which the departmental composition can change - new hires and academics leaving a department - to estimate the effect of performance pay on matching assortativeness in academia. If matching becomes more positively assortative after the reform, the difference in the productivity of new hires from before to after the reform should be larger for high quality departments. Furthermore, relatively lower productivity academics should be leaving high quality departments after the reform. Both predictions are borne out by the data and provide evidence that matching becomes more positively assortative after the reform. Moreover, this increase is larger in fields in which complementarities are stronger, which is in line with performance pay increasing positive assortative matching as driven by spillover effects. I find that the increase in positive assortative matching is large; amounting to a two- to threefold increase. The effect is driven by junior hires matching more positive assortatively, particularly so in high complementarity fields. This result aligns with the findings in Waldinger (2012) and (2010) that early-career academics experience the largest spillover effects.

A test for pre-existing trends finds no evidence of differences in trends between high quality and low quality departments in fields with different levels of complementarities. Furthermore, I test whether systematic differences in the hiring budget of high quality departments in fields in which complementarities are stronger could explain the results. In particular, if high quality departments in strong complementarity fields would have systematically larger hiring budgets, they could attract higher productivity academics by offering larger bonuses, and not because new hires benefit more from having highly productive colleagues under performance pay when spillovers are larger. Reassuringly, while I do find that a larger hiring budget enables departments to attract higher productivity academics after the reform, this does not affect the estimate of the effect of performance pay on matching assortativeness.

The rest of this thesis is organised as follows. Chapter 2 contains a description of the institutional background, with details on the pay reform in German academia, as well as the German academic system more generally. In Chapter 3, I provide details on the data set I use and its construction. Chapter 4 contains the analysis of the effort and selection effect of performance pay in academia, starting out with a theoretical framework, followed by the empirical analysis. Chapter 5 concerns the effect of performance pay on matching assortativeness. Here too, I first provide a theoretical model before moving on to the empirical analyses. Chapter 6 concludes.

2 Institutional Background

2.1 The Academic Pay Reform

In February 2002 a law introducing a new professorial pay scheme comprising a basic wage plus performance related bonuses was passed by Germany's parliament. States had until 1 January 2005 to implement the reform within their respective jurisdiction - only Bremen, Niedersachsen and Rheinland-Pfalz introduced it before this deadline (Detmer and Preissler 2005). The new "W-pay" scheme replaced the old "C-pay" scheme in which professorial salaries increased with age. The basic wage of the W-pay scheme is lower than C-wages for all but the lowest ages (Hochschullehrerbund 2009, Oeffentlicher-Dienst 2004), but the total pay under the W-scheme can exceed that under the C-scheme if an academic is paid large performance bonuses. Any contract for a professorial position entered (or renegotiated) after implementation of the reform falls under the W-pay scheme.

2.1.1 Performance Pay (W-Pay)

Under the new pay scheme, performance bonuses can be paid on three grounds: as wage supplements to attract outside professors or prevent professors from wandering off; as on-the-job bonuses for research or educational performance; and as supplements for professors taking on management tasks or roles (BMBF 2002). The first kind of bonus, the attraction or retention bonus, is paid to attract a professor or prevent him from leaving. These bonuses are generally awarded on the basis of a professor's qualifications and past achievements and performance (Detmer and Preissler 2005). Attraction or retention bonuses can be awarded permanently, but many states also allow for the option of awarding them for a fixed term (initially) or even as a one-off payment (Detmer and Preissler 2004, 2005). State laws and university statutes generally do not stipulate a maximum for attraction or retention bonuses.

The second class of bonuses are bonuses paid for on-the-job performance, awarded for accomplishments in research, teaching, art, mentoring and supervision (BMBF 2002). Research performance for instance may be demonstrated through the number and kind of publications, research prizes, patents and the award of external research funds, while e.g. exceptional teaching evaluations can serve to demonstrate special teaching achievements (Detmer and Preissler 2005).

Most universities pay the on-the-job bonuses for research and education through the so-called "Stufenmodell"; a system of performance levels each of which is associated with a bonus (Lünstroth 2011). Promotions to a higher level and the associated bonus are not necessarily granted permanently - they can be granted for a limited time period and renewed upon positive evaluation of performance. Most universities announce at the beginning of a year either both the number of levels and associated bonus pay or just the number of promotions to higher levels (in which case the associated bonus pay is generally laid down in the university's statutes) to be awarded in that year (Lünstroth 2011). The distribution of on-the-job bonuses through the Stufenmodell is therefore much like a promotion tournament (Harbring et al. 2004, Kräkel 2006, Lünstroth 2011). Both the number of levels and the associated pay varies greatly across universities; the number ranging from 2 (e.g. Augsburg and Erfurt University) to 10 (University of Trier), and the associated pay from 90 (Technical University of Berlin) to 2500 Euros per month (e.g. Bielefeld and Bremen Universities) (Lünstroth 2011).

Some universities pay on-the-job bonuses through a relative performance pay system ("Leistungspunkte Modell"). In this system, academics get awarded points for achievements in research and education and the university announces at the beginning of the year how much money will be available for on the job performance pay that year. Each academic then receives a share of the "prize pot" that is equal to his relative performance that year, making this essentially a Japanese style (J-type) tournament (Kräkel 2003, Lünstroth 2011).

The third kind of bonus takes the form of supplements that can be paid to professors for taking on management tasks or roles (BMBF 2002). These bonuses generally range between 200 and 600 Euros (for the dean) per month and are paid for the duration of the task or role only.

The reform also introduced the option for professors to be paid a supplement from third-party awarded funds for research or teaching projects for the duration of such projects (BMBF 2002)[1]. Some states stipulate that these supplements should not amount to more than the basic W-wage of the professor (Detmer and Preissler 2005).

Under the W-pay scheme, only tenured professors can be awarded bonuses - junior professors can only be paid a (non-pensionable) supplement of 260 Euros per month upon positive evaluation (Detmer and Preissler 2005)[2]. This is only a very small bonus compared to the total amount of bonuses that tenured professors may be awarded, which can go up to 5241,48 Euros[3], or even more in special circumstances[4](BMBF 2002, Detmer and Preissler 2005). Importantly, universities have discretion on how to award bonuses [5].

The academic pay reform includes a cost-neutrality clause that stipulates that the average professorial pay at the federal ("Bund") and state ("Land") level remain at the respective levels before the reform, so as to prevent the reform leading to either cost-cutting or a cost explosion (Handel 2005). The law does allow for states to increase their target level to, at most, the highest state average, as well as year-on-year increases of on average 2% (up to 10% in total) (BMBF 2002). Given that the base wage of the performance pay system is lower than most of the salaries under the age-related pay system, the cost neutrality requirement guarantees that the difference between C-salaries and W-base pay is paid as bonuses under the W-pay scheme. Handel (2005) calculates that with a mean professorial pay average of 71.000 Euros at universities, about 26% of this is available for performance pay bonuses for university professors[6]. In many states, the state's ministry of education implements the cost-neutrality requirement by calculating university-specific professorial pay averages that are to be used as guideline professorial pay average at each university (Handel 2005). The fact that on-the-job bonuses are distributed through, what are essentially tournament schemes, where the number and amount of bonuses to be won in a given year are announced at the beginning of a year suggests that the benchmark pro-

[1] These supplements were intended to motivate professors to take on activities as part of their academic job that they may have otherwise performed on the side (Handel 2005)

[2] Plus, in special cases an extra supplement per month not exceeding 10% of the basic W1 wage (Detmer and Preissler 2005)

[3] This limit is set at the difference between the basic wage of W3 and B10 (another, non-professorial pay scheme), which was 5241.48 on 1 August 2004 (Detmer and Preissler 2005)

[4] If the academic already earns bonuses that exceed this limit and a higher bonus is necessary to attract the academic to another German university or prevent him from wandering off to another German university (BMBF 2002).

[5] See Handel (2005) for a comprehensive overview of how much discretion higher education institutes have regarding hiring and pay decision after the reform in the different German states.

[6] For this calculation, Handel (2005) uses 2001/2002 data and assumes that the ratio of W2 to W3 professors at universities will be about the same as that of C3 to C4, namely 46:54.

fessorial pay average is reasonably binding each year (Harbring et al. 2004, Kräkel 2006, Lünstroth 2011).

2.1.2 Age-Related Pay (C-Pay)

In the old, age-related pay scheme, the monthly salary of professors increases every two years by roughly 170 Euros[7], from the age of 21 to the age of 49 (Hochschullehrerbund 2009, Oeffentlicher-Dienst 2004, Expertenkommission 2000). In contrast, the basic wage in the performance pay scheme does not vary with age, and the level is such that professors earn a higher before-bonus wage in the performance pay scheme at first, but once they get older, they would earn a higher basic wage in the age-related pay system. Depending on the specific pay level of an academic (C3, or C4 in the age-related system; W2 or W3 in the performance pay system), the crossing point of the basic wage schedules occurs at age 33 or 43 (Oeffentlicher-Dienst 2004, Handel 2005).

Before the pay reform, professors in the highest pay level of the age-related pay scheme (C4) could earn bonuses when they received offers after their first appointment as C4-professor. These bonuses were standardised to be around 650 Euros for the second C4-offer, and about 730 Euros for the third C4-offer from another university, and roughly 75% of this if a counter-offer of the home university was accepted (Detmer and Preissler 2006, Preissler 2006, Dilger 2013). By comparison, the average attraction and retention bonus in the W-pay system had already grown to 1187 Euros in 2006, and the average on-the-job performance to 1649 Euros (BMI 2007). Furthermore, only a small fraction of professors qualified for and received bonuses under the age-related pay system. Handel (2005) for instance calculates, using data from the Ministry of Science and Culture in Niedersachsen, that only about 16.5% of professors received attraction or retention bonuses in the age-related pay system. In contrast, any tenured professor in the performance pay system can receive bonuses, and already in 2006 about 77% of professors in the performance pay scheme did receive bonuses in the performance pay system (BMI 2007). Consequently, only about 3.55% of the total professorial pay volume was spent on attraction and retention bonuses in the age-related system, before the reform (Handel 2005, using data from Expertenkommission 2000), while an estimated 26% of the professorial pay volume was available for performance bonuses under the performance pay scheme immediately after

[7] Using pay tables valid as of August 2004 (Hochschullehrerbund 2009)

the reform(Handel 2005). Combined with the fact that, at most ages, the basic wage is lower in the performance pay system than in the age-related system, this means that a larger portion of professorial pay depends on performance and there is a greater spread in professorial pay in the W-pay system. The W-pay system therefore offers higher powered performance incentives than the old, age-related pay system.

2.2 Professorships

In order to qualify for a professorship, aspiring academics have to complete a PhD, as well as, traditionally, a post-doctoral qualification ("habilitation"). The habilitation involves working as part of the research group of a full professor, and is completed with a postdoctoral thesis (Fitzenberger and Schulze 2014, Pritchard 2006). In 2002 the German equivalent of assistant professorships ("Juniorprofessur") was introduced to supersede the habilitation (Pritchard 2006). Junior professorships can last up to six years and grant aspiring academics more independence than the habilitation (Fitzenberger and Schulze 2014). There are two tenured professorial ranks in Germany; the "ausserordentliche (or a.o.) Professur", which is the equivalent of an associate professorship, and the "ordentliche (o.) Professur", which is the equivalent of a full professorship (*Research and Academic Jobs in Germany* (2011)).

When a professorial position needs to be filled, an appointment commission is formed consisting of faculty professors, students and academic staff. The appointment commission compiles a top-3 list of candidates that, in turn, needs to be approved by a departmental committee (also consisting of professors, students, academic and non-academic staff). Professors always make up the majority in such commissions. The top-3 list of candidates is given to the state's ministry of higher education, which then decides which of the candidates on the list to appoint (Lünstroth 2011).

2.3 Higher Education Institutions

There are currently 397 higher education institutions in Germany that are either public or private but recognised by the state (*Hochschulkompass* 2014). The two main categories of higher education institutions are the universities ("Universitaeten") and the universities of applied sciences ("Fachhochschulen"). The former are more research oriented, the latter more vocationally oriented (Jongbloed 2009). There are currently 89 public universities in Germany, and I will focus on these institutions in this thesis (*Hochschulkompass* 2014)[8]. because the reform changes the pay schemes of academics at public higher education institutions only, and because German higher education institutions that are not universities, such as the universities of applied sciences ("Fachhochschulen"), are much more applied and faculty tends to publish much less (BMBF 2002). I would therefore not be able to use publication records to derive meaningful measures of productivity of academics at the latter institutions, and since I focus on the research output dimension of productivity, I will not consider them in this study.

Public universities are publicly funded, with most of the public funds coming from the respective state's ministry for higher education. Public subsidies make up around 80% of the income of a university, while additional research grants comprise roughly 15%. Of the additional research money, about a third is provided by private institutions (Kaiser et al. 2002). The public subsidies to universities have been traditionally subdivided in expenditure categories (line items) and personnel positions (described in the so-called "Stellenplan"), and determined in large part based on previous year's subsidies with only incremental changes. Any incremental changes to the budget would have to be negotiated with the state ministry. Recently, states have started to move towards more indicator-based budgeting, though the share of public funds allocated in this way is still small (up to 7%) (Jongbloed 2009). Public universities are thus greatly dependent on the state for their personnel and other expenditures and before the reform had little means or autonomy to pay professors a wage other than the age-related wage dictated by the old pay system.

[8]This was the number of public universities as reported by *Hochschulkompass* (2014) on 31 August 2014. The list of public universities that I consider for the empirical analysis is slightly different, i.a. because some institutions became universities only recently (e.g. the Hochschule Geisenheim University became a university on 1 January 2013 and the Hochschule für Film und Fernsehen Potsdam only in July 2014 (*Historie 1872 bis heute* 2014, *Filmuniversitaet* 2014)).

3 Data Description

For this project I constructed a new panel data set comprising the universe of German academics for the years 1993-2013. This individual level panel comprises 55132 academics who, at some point in the 15-year timespan, held a tenured position at a German public university[1]. The data set contains the affiliation, position (title) and whether the position is tenured in a given year, as well as the number of publications in that year, the number of years a person has been active in academia and the average number of publications in the preceding six years. Furthermore, the data set provides the year when postdoctoral qualifications were obtained and the year a person started working in academia. Finally, there is also some personal information such as gender, birth year and, if applicable, year of passing.

For the purposes of this research, I focus on academics who held a tenured position at a German public university between 1999 and 2013, because the reform changes the pay schemes of academics at public higher education institutions only, and performance bonuses can be earned in tenured positions only[2]. The reform was implemented for other German higher education institutions, such as the universities of applied sciences, too (BMBF 2002). However, as their name suggests, these institutions are much more applied and faculty tends to publish much less. I would therefore not be able to use publication records to derive meaningful measures of productivity of academics at these institutions, and given that I focus on research output and use publication records as a measure of this output, I do not consider higher education institutions other than universities in this study.

To construct the individual panel data set I draw from three main input data sets:

[1] The individual level panel is actually much bigger. It also contains academics at non-tenured positions, at other German institutions of higher education such as universities of applied science ("Fachhochschulen" in German), German private universities and higher education institutions in Austria and Switzerland.

[2] The German equivalent of assistant professors, "Juniorprofessoren" (awarded a W1-salary), can earn a very small yearly bonus only in the performance pay system.

Kuerschners Deutscher Gelehrten Kalender, Forschung & Lehre Magazine and ISI Web of Science. Kuerschners Deutscher Gelehrten Kalender (hereafter: DGK) is a comprehensive encyclopedia of academics affiliated with German universities (*Kuerschners Deutscher Gelehrten Kalender Online* 2013, *Kuerschners Deutscher Gelehrten Kalender* 2006, 2008). I use it as a register of the universe of academics affiliated with German universities and draw information regarding academics' personal information (full name, birth date, year of passing, gender) as well as professional information (academic affiliation at given points in time, start year of academic career in Germany, end year of academic career in Germany, self-reported information on career history) from it.

From the Forschung & Lehre Magazine (hereafter: FuL) I draw information regarding the timing of changes of affiliations and the obtainment of postdoctoral qualifications of these academics (*Forschung und Lehre* 1999-2013). FuL groups this information in 12 broad categories that encompass the field in which the respective academic is working. I classify academics registered in DGK under these same categories and define departments along the same lines. The FuL categories, and hence the departments that I distinguish, are: theology; philosophy and history; social sciences; philology and cultural studies; law; economics; mathematics, physics and computer science; biology, chemistry, earth sciences and pharmaceutics; engineering; agricultural sciences, nutrition and veterinary medicine; medicine (human); dentistry[3].

Finally, I use the ISI Web of Science database to compile publication records of the academics in my data set. Specifically, I draw the number of publications of an academic in a given year from the ISI Web of Science database for the years 1993-2013. I then weigh each publication by the two-year impact factor of the journal in which the publication appears. The impact factors are taken from the ISI journal citation report (JCR) of the year of publication[4].

I can match 83% of academics who appeared as having a tenured affiliation with a German university in FuL to academics listed in DGK on the basis of last name, initials and field. The 17% that I cannot match appears to be down to misspellings of names and erroneous affiliation changes information in FuL. 50% of changes in affiliations[5] are

[3]The department level panel consists of a total of 1068 departments, which amounts to an average of just under 11 departments per university.

[4]I have ISI JCR data for the years 2000-2013 only. I therefore use the average of the impact factors from JCR 2000 through JCR 2004 to weigh publications before 2000.

[5]Where at least one of the affiliations concerns a tenured position at a German university.

described in FuL, providing direct information on the timing of the change. Of the other half of changes, the year of change is given in the respective DGK record in 23% of the cases[6]. I infer the timing of the remaining affiliation changes from affiliation records of an academic at different points in time, the year they obtained postdoctoral qualifications as well as the start and end year of their academic career in Germany recorded in DGK. A detailed description of the construction of the individual level panel can be found in later sections in this chapter.

For the purposes of studying the effect of performance pay on matching assortativeness I derived a departmental level panel data set from the individual level data set. The departmental level data set comprises the departments of each of the 89 German public universities between 1999 and 2013. For each department, the panel contains the total number of tenured professors in a given year, the number of new hires into tenured positions, the number of academics already affiliated and in tenured positions, and the number of tenured professors that retire in a given year. Apart from the total number of new hires, the panel contains the number of new hires that start their first tenured position at a public university (junior hires) and the number of new hires that move from another tenured affiliation (senior hires). Furthermore, the panel also contains the number of people who leave a tenured position. For all of the categories of academics, the department panel contains average productivity variables, where the average productivity is calculated as the average number of publications weighted by impact factor, in the preceding six years. Lastly, the panel comprises data regarding the average productivity and the hiring budget of the departments, as well as the total number of retirees in a given year. The precise definition and construction of all these variables is described in the next section.

The rest of this chapter is organised as follows; I describe each of the three input data sets in more detail, before providing a description of the preparation, manipulation and matching procedures used to generate the eventual department-level panel data set used for the empirical tests in this thesis. All data handling was done using Python, unless indicated otherwise.

[6]This concerns self-reported career information. Because of the self-reported nature of this information and the corresponding self-selection bias this may introduce in my data set, I rely on FuL information regarding the timing of affiliation changes wherever I can. I checked for the consistency of the information on the timing of affiliation changes in FuL and DGK. The timing information in DGK differs from that in FuL for 5% of the individuals that change a (tenured) affiliation at least once.

3.1 Input Data sets

Kuerschners Deutscher Gelehrten Kalender (DGK) is a bibliographic and bibliometric encyclopedia of academics affiliated with German, Austrian and Swiss universities. All people who have passed the "venia legendi" and are both actively teaching and researching at a relevant university in Germany, Austria and Switzerland are included in DGK. The "venia legendi" encompasses the "habilitation" (a post-doctoral qualification that is acquired through publication of a habilitation thesis after up to six years of research as part of a full professor's research group ("Lehrstuhl")) and a qualification to teach at university level (the "Lehrbefugnis"). An exception to the venia legendi rule for inclusion in DGK are Honorarprofessoren and Juniorprofessoren. Universities considered relevant for DGK are generally those that can reward doctoral degrees ("Promotionsrecht"). This includes all public universities that I restrict attention to. Academics who move to a university outside of Germany, Austria or Switzerland are generally dropped from the encyclopedia, unless they personally request to remain included (Schniederjuergen 2013a). people who can no longer be verified to be affiliated with a university are dropped from the encyclopedia too. The information in DGK stems from academic calenders/teaching schedules, announcements of appointments by universities and in academic and professional journals, surveys, university websites, etc. (*Kuerschners Deutscher Gelehrten Kalender* 2006, 2008). De Gruyter Publishers, the current publishers of the DGK, have kindly supplied me with the editorial database underlying the online DGK edition (current up to 13 July 2013), as well as a copy of the exports from this database taken on 10-11-2006, 17-11-2008 and 27-09-2010. This database and its past exports contain the same information as the published DGK editions from the same years (that is, all records of people complying with the DGK inclusion criteria set out above), plus inactive records (of people who left (German, Austrian or Swiss) academia, passed away or could no longer be traced), activation dates of records (the date when a person first complied with the DGK inclusion criteria and was taken up in the database) and inactivation dates where applicable.[7]

[7] The first DGK edition to also be published electronically (as a CD-rom) was the 17th edition, the hard copy of which was released in 1996, the corresponding cd-rom in 1997. Subsequent CD-roms were released in 2007 and 2009, along with the corresponding 21st and 22nd editions of the hard copy DGK (*Kuerschners Deutscher Gelehrten Kalender Online* 2013). Since 2010 the DGK has an online version. The editorial database underlying this online version is updated continuously. The DGK editorial database was started in 1996, when the DGK data were migrated from the previous publisher to De Gruyter (Schniederjuergen 2013b). The earliest activation dates in the database however appear to be 1999.

Forschung und Lehre is Germany's largest higher education and research magazine that has been published monthly by the German higher education association (Deutscher Hochschulverband) since 1994 (*Forschung und Lehre, "Wir ueber uns"* 2014). Every magazine contains a section titled "Habilitation und Berufungen" with announcements of habilitations, the acquisition of the Lehrbefugnis, and the receipt, acceptance or rejection of academic (professorial) positions. These notifications are based on information from press releases from universities, newspapers and professional magazines as well as from readers/individual scientists (FuL 2002). Electronic copies of past Forschung und Lehre magazines from 1996 onwards can be downloaded from the "archive" section of the magazine's website (*Forschung und Lehre* 1999-2013). I use Forschung und Lehre magazines from 1999 to 2013 for the individual and department-level data sets for this thesis, so as to align with the years for which I have (activation) data from DGK.

The ISI Web of Science database (hereafter: ISI) is compiled and maintained by Thomson Reuters and can be accessed via the website apps.webofknowledge.com. From this database, I restrict attention to publications from the following databases: Science Citation Index Expanded (SCI-Expanded), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (AHCI), Conference Proceedings Citation Index - Science (CPCI-S) and the Conference Proceedings Citation Index - Social Sciences & Humanities (CPCI-SSH). I restrict the scope to publications with at least one of the authors with a German (work) address and published between 1993 and 2012, the records of which I downloaded from the ISI website.[8]

3.2 Preparation, Manipulation and Matching of Data

FuL provides information regarding the timing and specifics of the obtainment of habilitation and Lehrbefugnis as well as affiliation changes.[9] As for habilitation or Lehrbefugnis obtainment, I extract the name and current title of the person concerned, the current affili-

[8] I am currently working on a code that singles out academics who move to Germany from another country, so that I may compile more complete publication records for these academics on an individual level.

[9] I exploit the generally formulaic structure of the announcements in the "Habilitationen und Berufungen" section in FuL to distill the desired information regarding habilitationen and professorial offers in the Forschung und Lehre magazines from the text blocks in the magazine and put these in a tabular format. In the case of a habilitation and/or Lehrbefugnis announcement in FuL, the university at which the Habilitation and/or Lehrbefugnis was obtained is usually mentioned, as is the respective field. Professorial offer ("Berufung") announcements generally mention an academic's current university affiliation and title, the offer university and offered position (title and subject), as well as whether the offer was obtained ("erhalten"), accepted ("angenommen"), appointed ("ernannt") or rejected ("abgelehnt).

ation of the person and, if different, the university at which the qualification was obtained, the field in which the qualification was acquired, as well as the subject category under which the announcement was made in the FuL magazine. I take the month and year of the FuL issue in which the announcement was made to be the time when the qualification was obtained, backdated by four months to correct for the average printing lag. In the case of a professorial offer announcement, I record whether the offer was accepted, appointed or rejected, the name and current title of the person concerned, the current affiliation of the person, the offer university, offered position and field in which the position is offered, as well as the subject category under which the announcement was made in the FuL magazine.[10] Here too, I take the month and year of the FuL issue in which the announcement was made to be the time when the qualification was obtained, backdated by four months to correct for the average printing lag.[11]

I make the information in the three databases compatible by replacing university names in the FuL and DGK databases with unique identifiers, classifying all subject areas distinguished in DGK and ISI under 12 broad categories[12], mapping titles and positions to a unified dictionary of existing titles and positions, and classifying a title or position as being tenured or non-tenured[13]. Subsequently, I distill a list of unique academics from both the FuL and DGK records. In order to do so, I deduplicate the lists of academics from FuL and DGK on last name, initials and subject area[14].

[10] In case of multiple offers, I always record accepted or appointed offers first, followed by offers that are obtained. I record rejected offers last. In case of only obtained offers, I record offers from German universities first, otherwise the order is random.

[11] Offers that were only reported as being obtained by FuL are backdated by 2 months only, reflecting the fact that offer acceptance or rejection is reported two months later on average.

[12] These are the categories distinguished in the 'Habilitationen und Berufungen section' of FuL: theology; philosophy and history; social sciences; philology and cultural studies; law; economics; mathematics, physics and computer science; biology, chemistry, earth sciences and pharmaceutics; engineering; agricultural sciences, nutrition and veterinary medicine; medicine (human); dentistry.

[13] The following are tenured positions: C3-Professor/W2-Professor/Ausserordentliche Professor/Associate Professor and C4-Professor/W3-Professor/Ordentliche Professor/U(niversitaets)-Prof.

[14] DGK records each have a unique identifier in the underlying database.

3.2.1 Matching up Databases

I match academics appearing in FuL with academics in DGK on the basis of their last name[15], subject area[16] and initials[17]. Furthermore, I discard a potential match if:

a) a person's last (most recent) announcement in FuL is made while a potential match in DGK is over 67 years old[18] (based on birth year given in DGK)

b) a potential DGK match has a date of passing that falls before the last (most) recent announcement year in FuL

c) a potential DGK match is reported to have retired in DGK-year-x, while there are FuL announcements after year x

d) a potential DGK match is reported as having a tenured position before the habilitation year reported in FuL[19]

As mentioned above, 83% of academics who appeared as having a tenured affiliation with a German university in FuL can be matched to academics listed in DGK.

I match publications from ISI to academics appearing in FuL and DGK on last name, initials and subject area. If no match on last name, initials and field is possible, a match on last name and initials is attempted[20]. Whenever a match is found, I augment the publication count of a person in the given ISI publication year by the impact factor of the journal in which the publication appeared. I discard academics who share the same last name, initials and field are discarded from the data set, to prevent attributing publications of multiple different academics to multiple academics sharing last name, initials and field.

[15] Specifically; the name after the last space in the full name field, with potential hyphens of composite last names deleted (so e.g. Schmidt-Angel becomes SchmidtAngel).

[16] At least one of the FuL-field codes for the subject areas in which a person appears to work in DGK must be the same as the FuL-field code the person is classed under in FuL. If an academic does not have a subject area listed in DGK or if this subject area could not be classified under one of the FuL-field codes, a match is attempted on the basis of last name and initials only (but only if the subject area recorded in DGK could not be mapped to an FuL field code or if no subject area was recorded in DGK).

[17] Composite first names are separated first and the first letter of all name components are taken to be initials (e.g. Anna-Maria has initials A, M).

[18] German law mandates that academics retire at the age of 65 (Mohr 2007, *Bundesbeamtengesetz* 1985), so unless an academic moves abroad around the time of mandated retirement in Germany (cf. Mohr 2007), I should not observe FuL announcements regarding new affiliations for an academic who is past the age of retirement. I use 67 as cut-off instead of 65 to allow for some delay in a possible move abroad or FuL's reporting thereof.

[19] Where I allow for up to a one year lag in this announcement to allow for obtainment of a tenured position immediately upon passing the habilitation, as well as a publication lag in FuL.

[20] (but only if the subject area recorded in DGK could not be mapped to an FuL field code or if no subject area was recorded in DGK)

3.2.2 Creating an Individual-Level Panel

The starting point for the individual-level panel of affiliations over time are the FuL announcements. I supplement and check these with information from DGK. For any FuL offer announcement the current university of a person, his current position (title) and whether this concerns a tenured affiliation is filled back in time from the year before the FuL announcement year to the year that FuL reported as the year in which the person passed his habilitation or Lehrbefugnis; or, if this data is not available, the year in which the person is reported to have passed his habilitation in DGK[21]; or, if that is not available either, the activation year of the person's record in DGK, or the start year of the panel – whichever is earlier[22,23].

If the FuL announcement concerns an accepted offer, the new university, new position (title) and whether the position is tenured or not is filled forward from the year of the FuL announcement to the last year of the panel, or the year of passing or inactivation of the record if reported in DGK – whichever is earlier. If the FuL announcement concerns an appointment ("ernannt"), if a university different from the current university is given this is taken to be the offer university, if not, the current university is taken to be the offer university. The offer university, offered position (title) and whether this position is tenured or not is filled forward as in the case of an accepted offer. If the FuL announcement states that an offer was rejected, the current university, current position and whether the position is tenured or not is filled forward as above. Finally, for an announcement of a received offer ("erhalten"), the information regarding the offer university, position and whether the position is tenured or not is stored in a temporary list. If FuL reports the offer got accepted or rejected at a later date, the offer information is recorded in the manner described above for the respective announcement type. If there are no further FuL announcements regarding the offer, it is checked with information in DGK to attempt to confirm whether the offer was accepted or rejected. If the current university of the FuL announcement is matched by the university affiliation recorded in DGK, this affiliation is filled forward as above. If the

[21]This is based on self-reported information

[22]If an affiliation is already filled out in the year before the offer announcement, the current position is not filled backwards, but merely checked for consistency with the affiliation already recorded in the panel. If the two do not match up, an error message is created and the case is left for further, case-by-case evaluation.

[23]As mentioned before, the FuL announcement date is backdated by 4 months to correct for the average lag in reporting of offer information, so that the announcement of an accepted offer in e.g. February 2003 is interpreted as the offer being accepted in 2002.

offer university of the FuL "erhalten" announcement is matched by the university recorded in DGK, this affiliation is filled forward as above[24].

Whenever an academic changes affiliation according to data in DGK, but no announcement of a change appears in FuL, the start date of the new affiliation is taken from the self-reported career history in DGK, or, if that is not available, is taken to be the year after the previous DGK data year, or - if this is earlier - the minimum of the habilitation's year and activation year.

3.2.3 From Individual-Level to Department-Level Panel

For each academic, I derive 5 auxiliary variables. I define the 'start-date' of a person's academic career as the minimum of the year in which he first publishes, the year in which he received his habilitation or Lehrbefugnis as reported in FuL or DGK, and the activation date reported in DGK. I take the 'end-date' of a person's academic career to be the minimum of the last year in which I see a person publish, his date of passing and the inactivation date reported in DGK. Furthermore, for every year in the panel, I calculate the number of publication years, the sum total of impact factor weighted publications and the average number of impact factor weighted publications in the 6 years up to and including year t. In this time-span I only count years in which the person is considered academically active towards the number of publication years, and it is this number that I use as denominator of the average number of weighted publications. I then calculate the mean and standard deviation of the number of publication years, the sum of weighted publications and the average number of weighted publications at the department level, for the sample of academics who have a tenured affiliation at the university and are not retired yet. I use the year in which someone turns 66 as the year in which he no longer takes his seat due to retirement[25]. I also calculate the number of people retiring from a given department as well as from the university as a whole in a certain year, the number of new hires (defined as people not at department x in year t-1, or at x but not in a tenured position), affiliates (people in the same department, in a tenured position in year t-1) and the total faculty (sum of new hires and affiliates).

[24]If neither the current university nor the offer university from the FuL announcement is matched by information in DGK, a record of the mismatch is made in an error file and left for further, manual inspection. In the case of such a mismatch the current university, position and tenure variable are filled forward, as outlined above.

[25]The legal retirement age in Germany is 65 (Mohr 2007, *Bundesbeamtengesetz* 1985), so I take the year after this lustrum to be the year in which the pensioner's position may be refilled.

4 Effort and Selection Effects of Performance Pay in Knowledge Creation

4.1 Introduction

Universities constitute an important economic sector. The sector is not just important in terms of direct value and exports (McCormack et al. 2014), it also accounts for a significant portion of basic and applied research in an economy (Lach and Schankerman 2008). In turn, investments in research and education at universities have been shown to come with considerable localised spillovers (Kantor and Whalley 2014, Agrawal 2001) and act as a catalyst for innovation at a national level (Aghion et al. 2010, Acemoglu et al. 2006). Consequently, it is important to understand factors determining the performance of universities, ranging from general management practices (McCormack et al. 2014) to specific incentives like inventor royalty shares (Lach and Schankerman 2008 and 2004) and external factors such as a university's autonomy and the amount of competition it faces (Aghion et al. 2010). This chapter aims to add to this literature by zooming in to the level of academics and studying the effort and selection effect of performance pay in academia. Using the introduction of performance pay in German academia as a natural experiment, I find that performance pay increases effort by 35% and that more productive academics self-select into the pay scheme.

In order to estimate the effort effect of performance pay in academia, I exploit the fact that academics who make tenure just before the academic pay reform fall under the old, age-related pay scheme, while those who make tenure just after the reform are paid according to the new, performance pay scheme. If the timing of the tenure decision is exogenous, a differential change in productivity from before to after the reform between academics who make tenure just before the reform and those who tenure directly after the reform can be interpreted as the causal effect of performance pay on effort. I estimate this effort effect in a difference-in-differences framework using the academics who make tenure directly after the reform as treatment group and those who make tenure just before as control group. I find that performance pay increases effort by 35%. Roughly two-thirds (23%) of this effort effect can be ascribed to the tournament component of the performance pay scheme, with the remaining third (12%) caused by the piece-rate component.

I find no evidence of pre-existing trends, lending support to the identifying parallel trends assumption of the difference-in-differences estimation. Furthermore, I perform a number of tests to assess the validity of the assumption that the timing of the tenure decision is exogenous. I find no effect of the performance pay reform on academic effort in a placebo difference-in-differences, using two cohorts of academics who make tenure before the performance pay reform as placebo treatment and control. Since both these cohorts fall under the age-related pay scheme, the pay reform should not have any differential effect on their productivity. If, however, either lower productivity academics or higher productivity academics were able to speed up the tenure process, and if their productivity growth would slow down, respectively speed up after tenure, the placebo difference-in-differences would return a negative, respectively positive effort effect. I find no evidence of this.

The second part of the chapter is devoted to estimation of the selection effect of performance pay in academia. Academics who hold a tenured position when the reform is implemented switch from the age-related pay system to the performance pay system when they change their affiliation or position. Because pay no longer increases with age in the performance pay system but only with productivity, I expect more productive professors to be more likely to select into the performance pay scheme by changing position or affiliation. Hazard rate and survival function analyses confirm that more productive academics are indeed more likely to switch to the performance pay scheme. Moreover, an academic's productivity has a greater effect on his probability of selecting into performance pay if he is younger, in line with the performance pay scheme being relatively less attractive for older academics because of a larger difference in basic wage between the age-related and performance pay scheme.

As mentioned above, this chapter aims to contribute to the literature on university governance (cf. Haeck and Verboven 2012, Aghion et al. 2010, Belenzon and Schankerman

2009, Lach and Schankerman 2008, 2004) in particular, and to the literature on the organisation of knowledge creation (Grigoriou and Rothaermel 2014, Phelps et al. 2012, Jones 2009, Wuchty et al. 2007, Singh 2005, Audretsch and Feldman 2004, 1996, Jaffe et al. 1993)) more generally. By studying the effort and selection effects of performance pay in academia, this chapter studies one of the key components of human resource management, incentive systems, (cf. Lazear and Oyer 2012, Bloom and Reenen 2011, Oyer and Schaefer 2011, Lazear and Shaw 2007) in relation to university governance, thus building on and adding to the vast literature on incentives in organizations as well as that on university management and the organisation of knowledge creation.

Within the body of literature on incentives, numerous papers have examined how incentives affect worker productivity empirically, and many of those report significant positive effort effects of higher powered incentive schemes, both in the field (e.g. Muralidharan and Sundararaman 2011, Shi 2010, Lavy 2009, Bandiera et al. 2005, Shearer 2004, Lazear 2000a) and in the lab (i.a. Hossain and List 2012, Boly 2011, Dohmen and Falk 2011, Carpenter et al. 2010, Freeman and Gelber 2010, Bellemare et al. 2010, Ariely et al. 2009, Dickinson and Villeval 2008, Dickinson 1999). The kinds of higher-powered incentive schemes that have been shown to have positive effort effects are several; from piece-rate pay (cf. Dohmen and Falk 2011, Shi 2010, Bellemare et al. 2010, Ariely et al. 2009, Bandiera et al. 2005, Shearer 2004, Lazear 2000a, Dickinson 1999) or bonus pay (i.a. Hossain and List 2012, Muralidharan and Sundararaman 2011, Lavy 2009) to tournament schemes (cf. Carpenter et al. 2010, Freeman and Gelber 2010, Harbring et al. 2004) and monitoring regimes (e.g. Boly 2011, Dickinson and Villeval 2008). It is not a given that performance pay schemes would increase academic effort too, since academics are thought to be intrinsically motivated (McCormack et al. 2014), and extrinsic incentives might crowd out this intrinsic motivation (Dickinson and Villeval 2008, Besley and Ghatak 2005). The empirical results in this paper however provide evidence of a positive effort effect of both a piece-rate component and a tournament component of performance pay in academia. Though some papers have studied the effort effect of performance pay in education, either by estimating the effect of teacher incentives (cf. Muralidharan and Sundararaman 2011, Glewwe et al. 2010) or student incentives (i.a. Bettinger 2012, Leuven et al. 2011, Angrist and Lavy 2009, Angrist et al. 2009), this chapter is, to the best of my knowledge, one of only few works to study the effort effect of performance pay on academics, particularly with respect to research productivity. Lach and Schankerman (2008) and (2004) also study the effect of incentives on research productivity, but research productivity is measured at the university, not individual academic level in these papers, so that they cannot estimate the effort effect directly.

Within the body of literature on incentive schemes, there is also a number of papers studying sorting into pay schemes. Dohmen and Falk (2011) and Lazear (2000a) for instance study the selection effect of piece-rate schemes for workers in a lab experiment and in the field for windshield installers, respectively, and both papers find that higher productivity workers self-select into the higher-powered pay scheme. Leuven et al. (2011) report a similar finding for selection in tournament schemes for students. This chapter provides evidence that performance pay also has a positive and significant selection effect in academia. The selection effect thus reinforces the effort effect, both contributing to a greater productivity under performance pay.

This chapter is structured as follows: the next section presents a simple theoretical framework. The empirical analysis and core of the chapter make up section 3, with the first part focusing on the effort effect and the second part on the selection effect of performance pay. Section 4 concludes.

4.2 Theoretical Framework

In this section I present a simple theoretical model to analyse the effects of performance pay on the effort and selection of workers, modelled specifically to reflect the specific features of the performance pay reform in German academia, much like Lazear (2000a). As in Kräkel and Sliwka (2004), I consider a multi-stage game. At the beginning of every period, academics first select a pay scheme - either age-related or performance pay - and subsequently choose an effort level. They get paid at the end of a period. I solve for pure strategy subgame perfect Nash equilibria.

Suppose academics live and work for two periods. In each period, an academic's output depends on his effort e, ability $\theta \geq \underline{\theta} > 0$ and a random noise draw ε. The noise terms are assumed to be iid draws from a distribution $g\left(.\right)$ with mean zero and variance σ^2. In any given period, output is then given by:

$$y\left(e_i;\theta_i\right) = f\left(e_i;\theta_i\right) + \varepsilon_i \tag{4.2.1}$$

The production function $f(.;.)$ increases in both it's arguments, so $f_1, f_2 > 0$. Furthermore, let $f(0,.) = f(.,0) = 0$, $f_{11} \leq 0$ and $f_{12} > 0$. That is, I assume that no output is produced when an academic does not expend any effort, there are diminishing marginal returns to effort, and more able academics have a larger marginal productivity of effort at any effort level. Academics learn their ability at the beginning of the first period.

4.2.1 Baseline - Age-Related Pay

Before the pay reform, academics receive wages that increase with age. Though these age-related wages do not vary with an academic's performance, in what follows I allow for intrinsic motivation, which links an academic's utility to his performance[1]. Let academics be risk-neutral, expected pay-off maximisers. In the age-related system utility depends on the age-specific wage w_a, increases with output y_i at rate $r^a > 0$ (capturing intrinsic motivation and possible bonuses for C4-professors)[2], and decreases with the cost of effort $C(e_i)$. In any given period, the expected utility in the age-related pay scheme for academic i with age a who exerts effort e_i can then be written as:

$$E\left[U_{i,a}\right] = w_a + r^a E\left[y\left(e_i;\theta_i\right)\right] - C\left(e_i\right) = w_a + r^a f\left(e_i;\theta_i\right) - C\left(e_i\right) \tag{4.2.2}$$

The cost function $C(e_i)$ is convex in effort, so $C', C" > 0$, and $C(0) = 0$.[3]

[1] Furthermore, professors at the highest pay level in the age-related system, C4 professors, received bonuses for offers received after their first appointment as C4 professor (Dilger 2013, Detmer and Preissler 2006). Hence for C4 professors, part of their pay might depend on their (past) performance (to the extent that good performance increases the chances of receiving a second or third C4 offer). However, as noted in the previous chapter, these bonuses made up only about 3.55% of the total nationwide pay volume for professors according to Handel (2005), using data from Expertenkommission (2000). Furthermore, using data from the Ministry of Science and Culture of Niedersachsen, Handel (2005) finds that only about 16.5% of the Professors in Niedersachsen enjoyed such attraction bonuses.

[2] Here, the superscript "a" denotes age-related pay.

[3] Many papers model heterogeneity in agents through different cost functions, with higher ability agents having a lower marginal cost of effort (Lazear 2000b, Schotter and Weigelt 1992, Bull et al. 1987). As noted by Chen (2003), a specification in which ability affects productivity allows for an unambiguous definition of absolute and comparative ability and comparative statics with respect to ability. Heterogeneity is modelled as ability affecting productivity in i.a. Gürtler and Kräkel (2010), Kräkel and Sliwka (2004), Chen (2003).

4.2.2 Performance Pay

In the performance pay scheme, an academic's basic wage does no longer depend on his age, but any tenured professor can earn attraction or retention bonuses. A better performance increases the chances that an academic receives an offer and an attraction or retention bonus, effectively increasing the rate at which an academic's utility increases with his output. This can be modelled as the rate r being greater in the performance pay scheme: $r^p > r^a$ (where p denotes performance pay and a age-related pay)[4]. The performance pay scheme also introduces on-the-job performance bonuses distributed through tournaments. In a simple two-player tournament, academic i wins the tournament if $y_i > y_j$. This happens with probability

$$Pr\left(f\left(e_i;\theta_i\right)+\varepsilon_i > f\left(e_j;\theta_j\right)+\varepsilon_j\right) = G_{\varepsilon_j-\varepsilon_i}\left(f\left(e_i;\theta_i\right)-f\left(e_j;\theta_j\right)\right) \tag{4.2.3}$$

where $G_{\varepsilon_j-\varepsilon_i}(.)$ is the cdf of $\varepsilon_j-\varepsilon_i$ and $\varepsilon_j-\varepsilon_i \sim g\left(\varepsilon_j-\varepsilon_i\right)$. The distribution $g\left(\varepsilon_j-\varepsilon_i\right)$ has mean zero and variance $2\sigma^2$ because ε_i and ε_j are iid[5]. Denote the on-the-job performance bonus that can be won in the tournament by $b > 0$ and define $\triangle f_{ij}\left(e;\theta\right) \equiv f\left(e_i;\theta_i\right) - f\left(e_j;\theta_j\right)$. The expected utility under performance pay can then be written as:

$$E\left[U_i\right] = w + r^p f\left(e_i;\theta_i\right) + bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e;\theta\right)\right) - C\left(e_i\right) \tag{4.2.4}$$

In the first period, the basic wage an academic would earn in the performance pay system, w_1^p, is larger than the basic wage he earns in the age-related pay system, w_1^a, with probability λ. In period 2, the basic wage that academics earn in the age-related pay system (w_2^a) is larger than the basic wage they would earn in the performance pay scheme (w_2^p). This reflects the fact that the basic wage in the age-related pay scheme increases with age and the basic wage in the age-related pay scheme starts to exceed that in the performance pay scheme at age 33 or 43[6] (Oeffentlicher-Dienst 2004, Handel 2005).

Future utility is discounted by a factor $\delta, 0 < \delta < 1$.

[4] I thus assume that if the extrinsic motivation provided by the attraction and retention bonus crowds out intrinsic motivation, this crowding out is less than one-for-one. In the empirical section, I will provide evidence that is in line with this assumption.

[5] This follows the specification of the winning probabilities in a rank order tournament in Lazear and Rosen (1981).

[6] The age at which the basic wage in the age-related pay scheme first exceeds that in the performance pay scheme is 43 only if a C3-professor is offered a W3-professorship.

4.2.3 Effort Effect

I solve for subgame-perfect equilibria of this two-period, two-stage game by backward induction. Hence, I first solve for the optimal effort level in the second period. Under the age-related pay scheme, academics maximise 4.2.2 with respect to effort e_i. The optimal effort level e_i^* is then given by the first order condition:

$$r^a \frac{\partial f\left(e_i^*;\theta_i\right)}{\partial e_i} = C'\left(e_i^*\right) \tag{4.2.5}$$

By the implicit function theorem, we have that

$$\frac{de}{d\theta}\left(e_i^*\right) = \frac{-r^a f_{12}\left(e_i^*,\theta\right)}{r^a f_{11}\left(e_i^*,\theta\right) - C\text{"}\left(e_i^*\right)} > 0 \tag{4.2.6}$$

Hence higher ability academics exert greater effort in equilibrium even in the age-related pay system, provided academics are intrinsically motivated. It follows that higher ability academics also produce more output in equilibrium.

In the performance pay system, academics choose their effort level to maximise 4.2.4. As I solve for pure strategy Nash equilibria, academic i takes e_j as given when determining his effort level, and similarly for j (Lazear and Rosen 1981). If an interior solution exists, i's equilibrium effort under the performance pay scheme is given by the first-order condition:

$$\left[r^p + bg_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e^*;\theta\right)\right)\right] \frac{\partial f\left(e_i^*;\theta_i\right)}{\partial e_i} = C'\left(e_i^*\right) \tag{4.2.7}$$

A similar first-order condition gives j's equilibrium effort.

An interior solution exists if σ^2 is sufficiently large, so that the rate of change of the probability density $g\left(\varepsilon_j - \varepsilon_i\right)$ is not too large positive (Bhattacharya and Guasch 1988, Lazear and Rosen 1981). More precisely, an interior solution exists if the following second-order condition is met:

$$\left[r^p + bg_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e^*;\theta\right)\right)\right] \frac{\partial^2 f\left(e_i^*;\theta_i\right)}{\partial e_i^2} + bg'_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e^*;\theta\right)\right)\left(\frac{\partial f\left(e_i^*;\theta_i\right)}{\partial e_i}\right)^2 - C''\left(e_i^*\right) < 0 \tag{4.2.8}$$

Proposition 1 - Effort Effect: *Academics exert greater effort in any given period in the performance pay scheme than in the age-related pay scheme in a pure strategy subgame perfect Nash equilibrium.* (Proof in Appendix A)

Corollary 1: *Academics produce more output under the performance pay scheme.*

Corollary 1 follows immediately from Proposition 1 and the assumption that output strictly increases with effort.

Because both the attraction and retention bonuses and the on-the-job performance tournaments increase the marginal return to productivity[7], academics are expected to exert greater effort and hence produce more output under performance pay than under age-related pay. I would therefore expect the academic pay reform to have a positive effort effect.

4.2.4 Selection Effect

When the academic pay reform is implemented, academics have different ages. Equivalently, in the model an agent might be at the beginning of period one or two when the reform is implemented and he first gets to select a pay scheme. The basic wages of the two pay schemes compare differently at different ages, with the difference between age-related basic wage and the basic wage in the performance pay scheme being larger positive for older academics. The performance pay scheme is therefore relatively more attractive in the short-run for younger academics, who are at the beginning of period 1 when the reform is implemented, than for older academics, who are at the beginning of period 2 when they first get to choose a pay scheme.

Proposition 2 - Selection Effect: *If at least some academics prefer to remain in the age-related pay scheme after the pay reform, only higher ability academics, whose ability θ exceeds a threshold value θ_1^0 prefer to select into performance pay. This threshold ability level is larger for older academics.* (Proof in Appendix A)

When comparing utility under age-related pay with that under performance pay, there are two counteracting effects. On the one hand, the portion of utility that depends on output and, by extension, effort, is larger in the performance pay scheme than in the age-related pay scheme for any effort level. On the other hand, the basic wage is larger in the age-related pay scheme for most ages. Because higher ability academics have a greater marginal productivity, and because the return to productivity in the performance pay scheme is larger, the former effect is more likely to outweigh the latter for higher ability

[7] Provided extrinsic motivation crowds out intrinsic motivation less than one-for-one.

academics. Furthermore, since the difference between the age-related basic wage and the basic wage in the performance pay scheme is larger for older academics, the drop in basic wage is more likely to outweigh the higher return to productivity under performance pay for older academics. Put another way, older academics need to be of higher ability in order for the higher return to productivity to outweigh the drop in basic wage when switching to performance pay.

The next section reports the results of empirical tests of the hypotheses regarding the effort and selection effect of performance pay in academia put forward in propositions 1 and 2.

4.3 Empirical Analysis

In order to study effort and selection effects of performance pay in knowledge creation, I use the individual level panel data set that I constructed for this research project and that encompasses the affiliations and productivity measures of the universe of academics in German academia for the years 1999-2013. This individual level panel comprises 55132 academics who held a tenured position at a German public university at some point between 1999 and 2013. The data set provides information regarding an academic's affiliation, reporting whether his position is tenured, and whether he is affiliated with a public university in a given year. Furthermore, the data set contains a variable for the impact factor-weighted[8] number of publications of an academic in a given year and the average number of weighted publications in the previous six years. Furthermore, the data set provides the year in which an academic obtained his postdoctoral qualification, as well as the year a person started working in academia. Finally, there is a birth year variable and, if applicable, year of passing.

As discussed in the previous chapter, I restrict attention to academics who held a tenured position at a German public university between 1999 and 2013, because the reform only changes the pay schemes of academics at public higher education institutions, and performance bonuses can be earned in tenured positions only[9]. I discard higher education institutions other than universities, because I focus on research output, and the research

[8] A publication is weighted by the impact factor of the journal in the publication year, where the impact factors are taken from *Journal Citation Report* (2000-2012).

[9] "Juniorprofessors", the German equivalent of assistant professors, can earn a yearly bonus in the performance pay system, but this is only very small compared to the bonuses tenured professors can earn.

output of universities is incomparable to that of other higher education institutions.

4.3.1 Descriptive Statistics

Table 4.1, Panel A reports a few summary statistics for the individual level panel used for the effort effect analysis below. The main variable of interest is the weighted number of publications of academics in a given year. As the summary statistics in the table show, academics at public universities are on average more productive than academics at other higher education institutions (compare rows 1 and 2), while tenured professors at public universities are more productive than nontenured academics at public universities (compare rows 2 and 3). Academics who are in the early stages of their academic career and manage to obtain a tenured position at a public university are more productive still (cf. row 4). Row 4 shows the weighted number of publications of academics who make tenure at a public university either just before the reform, in 2003 or 2004, or directly after, in 2005 or 2006[10]. The former cohort comprises 2193 academics, the latter 1524, and I restrict attention to these cohorts to estimate the effort effect of performance pay as explained below.

4.3.2 Effort Effect

In order to identify the pure effort effect of the introduction of performance pay in German academia, I use the fact that any contract for a professorial position at a public university in Germany signed or renegotiated as of 1 January 2005 necessarily falls under the performance pay scheme, whereas any contract signed before this date falls under the old, age-related pay scheme[11]. Accordingly, academics who start their first tenured affiliation before 2005 continue to fall under the age-related pay scheme[12], whereas academics who

[10]To be exact, I restrict attention to academics who make tenure in 2003 or 2004, respectively 2005 or 2006, and who are affiliated with a public university at some point in the post-reform period. This also includes academics who spent some of the post-reform period at another higher education institution. As the summary statistics show, academics at higher education institutions other than universities publish less on average, and if they are at a private university, they are not affected by the pay reform. The productivity of academics not at public universities should therefore change less in response to the reform, and the below estimates of the effort effect are a lower bound. The estimates of the effort effect reported below are robust to restricting attention to academics who were affiliated with public universities throughout the post-reform period, and, as expected, larger.

[11]With the exception of Bremen, Niedersachsen and Rheinland-Pfalz, who introduced performance pay before this deadline (in 2003 and 2004, respectively) (Detmer and Preissler 2005). Note that using 2005 as uniform before-after cut-off yields a conservative measure of the effort effect, since some of the control group is in fact already treated before this time.

[12]They would be promoted to a higher pay grade in the age-related pay scheme when making tenure (Detmer and Preissler 2004).

make tenure after 2004 switch to the performance pay scheme upon making tenure. If the timing of the tenure decision is exogenous, the performance incentives that first-time tenured affiliates face is exogenous as well. I can then identify the pure effort effect of performance pay on academic productivity by comparing the change in productivity of academics who start their first tenured affiliation before 2005 (the control group) with the change in productivity of academics who start their first tenured affiliation as of 1 January 2005 (the treatment group) from before to after the pay reform. However, it is not ex ante obvious whether the timing of the tenure decision is exogenous. In particular, academics could try to speed up the process in order to avoid the performance pay system. I provide a number of tests that yield no evidence of this below.

Table 4.2 shows the unconditional means of academic productivity for the treatment and control group and before and after reform periods separately[13]. I use a two-year window before and after the reform to define the treatment and control group in order to abstract from seniority effects. Thus the treatment group consists of academics who first made tenure in 2005 or 2006, while the the control group consists of academics who first made tenure in 2003 or 2004. Academic productivity is defined here as the number of impact-factor weighted publications of academic i in field f in year $t + x_f$, where x_f denotes the average publication lag in field f. The average publication lags are taken from Björk and Solomon (2013) and differ across fields. I have data for 6 years before the reform (1999-2004) and 9 years after the reform (2005-2013).

The difference in means estimate in Table 4.2, Panel A, column 3, row 1 shows that there is no significant difference in productivity between the treatment and control group before the reform. After the reform however, the average productivity in the treatment group is significantly larger than that in the control group (cf. Column 3, row 2). Moreover, the difference between these two differences is positive and significant (cf. column 3, row3). If assignment to the treatment and control group was indeed exogenous, so that, absent the reform, the treatment group's productivity would have followed the same trend as that of the control group, this difference-in-differences estimate is an estimate of the causal effect of performance pay on academic effort. This effort effect is economically large; amounting

[13]The means and differences in means are estimated through a pooled OLS regression of the lagged, weighted number of publications of academics on a constant, a treatment dummy, a post dummy (that is 1 for the years after the reform (2005-2013), and 0 before, so that I estimate persistent shifts in productivity from before to after the reform) and a treatment*post interaction term. I estimated standard errors, clustered by individual academic.

to a 35% increase in academic productivity relative to the pre-reform productivity in the control group. In what follows, I perform several checks of the validity of this estimate of the effort effect.

The differences in means in columns 1 and 2 of row 3 in Panel A provide a first such check. Reassuringly, the average productivity of both the control group and the treatment group increases from before to after the reform. Hence the positive difference-in-differences estimate in column 3, row 3 really comes from a greater increase in the average productivity of the treatment group than the control group instead of a (larger) decrease in productivity in the control group. The fact that the average productivity of the control group also increases after the reform is consistent with a positive average treatment effect. This aligns both with increased effort from academics who still fall under the age-related pay scheme to up their chances of receiving a lucrative job offer in the performance pay scheme, and with greater positive spillover effects through increased positive assortative matching (Ytsma 2015).

In order to test the identifying assumption that the assignment to treatment and control is exogenous, I report the results of a placebo difference-in-differences estimation in Table 4.2, Panel B. Here, the placebo treatment group consists of academics who first made tenure in '03/'04 (the control group in Panel A), and the placebo control group consists of academics who first made tenure in '01/'02. If academics were able to influence the tenure clock, I would expect lower productivity academics to try to move up the tenure decision if it would allow them to stay in the age-related pay system. If not just the level but also the growth rate of academic productivity is smaller for lower productivity academics, this would decrease the growth in average productivity of the control group used above (academics first tenured in '03/'04) from before to after the reform. Consequently, the change in average productivity of academics who first tenured just before the reform (in '03/'04) would then be smaller positive (or larger negative) than the change in average productivity of an earlier cohort (academics who first made tenure in '01/'02). As the placebo difference-in-differences estimate in column 3, row 3 shows however, I do not find evidence of such selection into the '03/'04 first tenure cohort. The average productivity of both the placebo treatment and control group increases from before to after the reform (cf. columns 1 and 2, row 3), and this increase is even slightly larger for the placebo treatment group than the placebo control group, though not significantly so.

4.3.2.1 Baseline Difference-in-Differences

As a next step I estimate the effort effect in a parametric difference-in-differences model. In particular, I estimate the following equation as a Fixed Effects model (at the individual level) in order to estimate the effort effect of the introduction of performance pay:

$$Y_{i,f,t-x_f} = \alpha_i + \beta_1 t + \beta_2 post'05 + \beta_3 post'05 * Treatment_i + u_{i,t} \qquad (4.3.1)$$

The corresponding estimation results are shown in Table 4.3. The dependent variable, $Y_{i,f,t-x_f}$, denotes the lagged number of impact factor weighted publications of academic i in field f in year $t - x_f$, where x_f denotes the average publication lag in field f as before. The variable $post'05$ is 1 as of 2005 and 0 beforehand. The $Treatment$ variable is 1 for academics who start their first tenured affiliation at a public university in 2005 or 2006, and 0 otherwise. I restrict the sample to include only those academics who start their first tenured affiliation at a public university in 2003, 2004, 2005 or 2006. The $post'05 * Treatment_i$ interaction term is therefore a difference-in-differences estimate of the effort effect of performance pay in knowledge creation, with academics who start their first tenured affiliation at a public university in 2003 or 2004 as the control group. Robust standard errors, clustered at the individual level are reported throughout.

Column 1a in Table 4.3 reports the results from the baseline regression without linear time trend t, column b shows the results from the baseline regression with linear time trend. The $post'05 * Treatment_i$ interaction term is always positive and significant, and in the same order of magnitude of the $post'05$ coefficient estimate. The estimate of the interaction term implies that the academics starting their first tenured affiliation under the performance pay system (the treatment group) produce on average about one and a half weighted publication more than academics who started their first tenured affiliation under the age-related pay system in every year after the reform. I thus find evidence of an effort effect that is highly significant and economically large.

The $post'05$ coefficient estimate in Column 1a is also positive and significant, and economically large. Including the time trend however turns the $post'05$ coefficient negative and significant. This could be caused by mean reversion after promotion (here: making tenure), as argued in Lazear (2004). The linear trend itself is positive and significant, providing evidence that the number of weighted publications increases over time throughout the entire sample period. All in all, this suggests academics produce more as they grow older (at least at the beginning of their academic career), and implies I find no evidence that the reform increases the productivity of the control group as well.

4.3.2.2 Pre-Existing Trends

To further validate the identification strategy, I test for pre-existing trends. It could be that the positive and significant $post'05 * Treatment_i$ interaction term in the baseline simply reflects pre-existing trend differences between the treatment and control group. As a first pass, I therefore estimate the following simple difference-in-differences regression equation including a full set of treatment dummy * year dummy interactions:

$$Y_{i,f,t-x_f} = \alpha_i + \sum_{yr=2001}^{2013} (\beta_{1,yr} yr + \beta_{2,yr} yr * Treatment_i) + u_{i,t} \tag{4.3.2}$$

Here, yr is a year dummy[14], all other variables are as specified above and all regressions contain individual fixed effects. As for the baseline regression 4.3.1, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006, so that - as before - $Treatment$ is 1 if an academic starts his first tenured affiliation in 2005 or 2006 and 0 if he starts his first tenured affiliation in 2003 or 2004 (the control group). The point estimates and 90% confidence intervals of the corresponding interaction terms are plotted in Figure 4.4.1[15]. These interaction terms give the year-on-year productivity differences between the treatment and control group.

The figure shows no sign of a pre-existing trend: the estimates of the year-on-year differences first become positive and significant in 2005 only. They subsequently remain positive and significant for most of the post-reform period, so for most of the post-reform period, the productivity of the treatment group is statistically significantly larger than that of the control group.

As an alternative test for pre-existing trends, I augment the baseline model in equation 4.3.1 with three placebo *post* variables - $post\_t-3$, $post\_t-2$, $post\_t-1$ - and their interactions with the $Treatment$ variable[16]. These placebo *post* variables are 1 as of

[14] Note that I include year dummies here (e.g. the 2005 dummy is 1 in 2005 only, and 0 otherwise) and not post variables as in the baseline model in equation 4.3.1 (the post variable is 1 as of 2005 and 0 before).

[15] The estimation results are also reported in column 1 of Table 4.7 in appendix B.

[16] Where, as in the baseline model (4.3.1), the $Treatment$ variable is 1 for academics who start their first tenured affiliation at a public university in 2005 or 2006, and 0 otherwise.

2002, 2003 and 2004, respectively, and 0 beforehand. The estimation results in Table 4.3, columns 2a and 2b confirm that there is no evidence of a consistent pre-existing trend, since the interactions of the placebo *post* variables with the $Treatment$ variable are not all positive and significant. Reassuringly, including the pre-trend terms does not render the $post * Treatment_i$ interaction term insignificant either, though the size of the coefficient does decrease.

Furthermore, I find evidence of an announcement effect in these regressions, because the $post\_t - 3 * Treatment_i$ interaction term is positive and significant. This means that academics who would tenure after the performance pay reform comes into effect (as of 2005) and hence would get paid under the performance pay scheme once tenured, step up their game the moment the reform is announced (in 2002), so as to up their chances of earning bonuses in the future[17]. The effect amounts to about half a weighted publication more per year since the announcement of the reform; a 12% increase in academic productivity relative to the unconditional pre-reform productivity in the control group[18]. Given that only the, effectively, piece-rate component of the performance pay regime[19] takes effect from the moment the reform is announced, this 12% increase in productivity can be interpreted as the effort effect of the piece-rate component only. This means that the effort effect of the tournament component of the performance pay scheme brings about a 23% increase in productivity. The effort effect of competitive (tournament) pay is thus almost twice as large as the effort effect of piece-rate pay.

4.3.2.3 Placebo Experiment

As a final set of tests of the identification strategy, I perform the same regressions for a placebo treatment and control group. In Table 4.4 columns 1a and 1b, I show the results of a placebo difference-in-differences baseline regression (equation 4.3.1) with the placebo treatment group comprising academics starting their first tenured affiliation in 2003 or 2004 (the control group in the baseline regression in Table 4.3) and the placebo control group made up of academics starting their first tenured affiliation in 2001 or 2002. Given that both these groups fall under the age-related pay system, there should be no differential

[17] Note that this evidence also aligns with the assumption that the extrinsic motivation provided by the attraction and retention bonuses does not (completely) crowd out intrinsic motivation, since the prospect of bonuses related to (past) performance causes academics who will tenure after the pay reform to increase their effort the moment the reform is announced.

[18] Cf. Table 4.2, column 2, row 1.

[19] Namely the prospect of attraction bonuses, which are linked to (past) performance.

effort effect once the reform gets implemented. The fact that the $post * Placebo\_Treat_i$ interaction term is not significant is in line with this. As discussed in the previous section, these results do not provide evidence of endogenous selection into treatment and control group either[20].

Figure 4.4.2b depicts the confidence intervals of the interaction terms in equation 4.3.2 when the sample is restricted to academics who started their first tenured affiliation at a German public university in 2001, 2002, 2003 or 2004, so that $Treatment$ is 1 if an academic starts his first tenured affiliation in 2003 or 2004 (the placebo treatment group) and 0 if he starts his first tenured affiliation in 2001 or 2002 (the placebo control group). In line with the results in Table 4.4, columns 1a and 1b, most of the year-on-year productivity differences between placebo treatment and placebo control are not significantly different from zero in the post reform (2005-2013) period, again allaying concerns about endogenous selection into the treatment group.

Figure 4.4.2b also shows that the three interaction terms in the years surrounding tenure of the treatment group (2003, 2004, 2005) are positive and significant, This is to be expected, since academics need a good publication record to make tenure, and will therefore make every effort to publish (more) when their tenure clock is running out. Taking into account that some publications counting towards tenure may not actually have been published but only accepted by the time the tenure decision is made, allows for a relative upswing in publications by newly tenured academics the year after they make tenure. Such a surge in productivity around the time academics first make tenure would however only bias the baseline difference-in-differences estimates of the effort effect (Table 4.3) if it causes a permanent increase in the level of productivity from the moment of first tenure. In order to test whether this is the case, I finally estimate the placebo difference-in-differences model augmented with three placebo $post$ variables - $post\_t-3$, $post\_t-2$, $post\_t-1$ - and their interactions with the $Placebo - Treatment$ variable. The corresponding estimation results in Table 4.4, columns 3 and 4 show that none of the interactions are positive and significant, thus I find no evidence of a permanent upswing in the productivity level from

[20]Specifically, if lower productivity academics were able to speed up their tenure clock and if their productivity would slow down more than the average mean reversion in the placebo control group upon making tenure, the $post\_t - 2 * Placebo\_Treat_i$ interaction term or $post\_t - 1 * Placebo\_Treat_i$ interaction term would be negative and significant. If, on the other hand, higher productivity academics were able to speed up their tenure clock and if their productivity would slow down less than the average mean reversion in the placebo control group, the $post\_t - 2 * Placebo\_Treat_i$ interaction term or $post\_t - 1 * Placebo\_Treat_i$ interaction term would be positive and significant. I do not find evidence of either type of selection.

the moment an academic makes tenure.

4.3.3 Selection Effect

Apart from academics who tenure after the reform, and therefore necessarily fall under the performance pay scheme, academics who already have a tenured affiliation before 2005 can also select into the performance pay scheme by changing affiliation or position, or by opting into the pay scheme while retaining the same position. I do not have information on the latter, though Detmer and Preissler (2005) report that only a small number of professors chose to opt into the W-pay scheme in their current position. I do however observe professors changing affiliation or position and, consequently, changing into the performance pay scheme. I exploit this information to analyse the selection effect of the reform. Given that a professor's pay increases with performance through bonuses in the performance pay scheme, but no longer with age as in the age-related pay system, I expect more productive academics to be more likely to select into the performance pay scheme[21]. I test this hypothesis through hazard rate and survival function analysis in this section.

For this purpose, I derive survival data from the individual panel data set, in which I focus on academics switching from the age-related pay scheme to the performance pay scheme by renegotiating the contract of their current tenured position[22] or by changing to another tenured position, possibly at another university. In order to abstract from academics entering the performance pay scheme because they make tenure after the reform, I restrict attention to academics who already hold a tenured affiliation before 2005. There are 37571 such academics and I observe a total of 3376 switches in a total of 248107 periods (years) that these academics can switch from the age-related to the performance pay system (cf. Table 4.1, Panel B).

Figure 4.4.2 shows the Epanechnikov kernel density estimates of the hazard function for switches from age-related to performance pay for academics whose average productivity falls in the top decile or bottom 90% of the average productivity distribution[23]. Panel a of

[21] Cf. proposition 2 in the theoretical model.

[22] I assume that, whenever an academic receives an offer, he either accepts and changes position, or rejects and renegotiates his current contract. In either case, the academic switches to the new performance pay scheme if the change or renegotiation happens after the reform. If there are academics who do not at least renegotiate their contract when they receive an offer, these academics are more likely to be of a lower productivity type, and including them in the pool of switchers would reduce the estimate of the selection effect I find.

[23] Because only academics who already had a tenured affiliation before the reform can switch from age-related to performance pay by changing affiliation or position, only these changes are considered here.

Figure 4.4.2 uses contemporaneous average productivity measures, while Panel b employs the average productivity data from 2005. Because an academic's average productivity is calculated as the average number of impact factor weighted publications in years t-6 through t-1, the average productivity measures in 2005 abstract from the effort effect of the performance pay reform. In both figures, the hazard rate for switching to the performance pay scheme is clearly greater for top decile academics throughout, so higher productivity academics are more likely to sort into the performance pay scheme. The Kaplan-Meier estimates of the survival functions of staying in the old, age-related scheme in Figure 4.4.3 show the same result. A log-rank test of the equality of the survival functions of top decile academics and bottom 90% academics rejects the equality of the survival functions at the 1% significance level[24].

I estimate the selection effect of the introduction of performance pay parametrically by estimating the following Weibull proportional hazard model:

$$\lambda_{i,t} = \rho * exp\left[\beta_0 + \beta_1 \bar{y}_{i,t} + \beta_2 age_{i,t} + u_{i,t}\right] * t^{\rho-1} \tag{4.3.3}$$

The results of the estimation of this Weibull model are presented in Table 4.5. The model estimates the hazard ratio for academics moving from one tenured affiliation to another tenured affiliation at a public university[25]. As of 2005, any such move implies switching from the age-related pay system to the performance pay system. In columns 1a and 2a of Table 4.5, I only use the average productivity of academic i in year t: $\bar{y}_{i,t}$ (the average weighted number of publications of academic i in years $t-6$ through $t-1$) as explanatory variable. In columns 1a and b, I use a contemporaneous measure of average productivity, while in columns 2a and 2b, an academic's average productivity in 2005 is used to abstract from effort effects. In columns 1b and 2b, I also include an age variable. This age variable is calculated using an academic's reported birth year whenever available[26], and equal to a synthetic age otherwise. I calculate synthetic birth years by subtracting the average age at habilitation or career start of academics for whom I do have a birth year[27] from the year of habilitation or career start of the academics for whom I do not have a birth year.

[24]When using the contemporaneous average productivity measure, the log-rank test returns a Chi-squared statistic of 167.38 (p-value 0.0000). The log-rank test returns a Chi-squared statistic of 101.15 (p-value 0.0000) when using the 2005 average productivity measure.

[25]While being younger than 66 and hence not retired.

[26]I have birth year data for 43.3% of academics in the data set. In order to prevent selection bias, I supplement the birth year data with synthetic birth year data in the analyses.

[27]The average age at habilitation or career start is 38.

Synthetic age is then equal to the age imputed using the synthetic birth year for academics for whom I do not know the actual birth year. I restrict the sample to academics with a tenured affiliation at a public university after the reform, since only these academics have the choice (i.e. are "at risk") of switching pay scheme.

The coefficient estimate of the average productivity of an academic, $\bar{y}_{i,t}$, is positive and significant throughout, suggesting that more productive academics are more likely to select into the performance pay scheme. The coefficient estimate in column 1a implies that one extra impact factor weighted publication on average increases the probability of selecting into the performance pay scheme by 0.3%. Adding age as covariate increases the size of this coefficient. This is not surprising, given the strong negative correlation between age and average productivity. Indeed, one extra year of age is associated with an almost 9% decrease in the probability that an academic will select into the performance pay scheme. Accordingly, after controlling for age, the Weibull parameter ρ changes from being a precisely estimated zero - indicating hazard does not change over time - to being positive and significant, in line with increasing hazard over time. This implies that academics are more likely to switch to the performance pay system, the longer this system has been around, perhaps because any uncertainty regarding the practical implications of the new pay scheme is reduced as time goes by. The results are robust to using a contemporaneous average productivity measure in columns 2a and 2b[28].

4.3.3.1 Heterogeneous Hazard Rates

I next explore whether the hazard rates and their relationship with an academic's productivity vary by age category, and in particular whether I find support for the prediction in proposition 2 that older academics need to be of relatively higher ability than young academics in order to prefer to switch to performance pay.

Table 4.6 reports the results of the separate estimation of the Weibull model given in equation 4.3.3 for academics in different age categories. The respective samples are restricted to observations of academics younger than 37 in columns 1a and 1b, between the ages of 37 and 48 in columns 2a and 2b, and 49 years of age or older in columns 3a and 3b. These categories are chosen so that the start age of the second category is the average age at which the age-related basic wage starts to exceed the basic wage

[28]Estimation of equivalent Cox proportional hazard models yield similar results (results available upon request)

under performance pay, while the start age of the third category coincides with the age at which academics earn the highest possible basic wage in the age-related system (and hence no longer experience age-related increases in this basic wage). The average productivity variable is a contemporaneous measure of average productivity in the a columns, while I use an academic's average productivity in 2005 in the b columns to abstract from effort effects.

As expected, the effect of productivity on the switching hazard rate is largest for academics in the youngest age bracket; one extra impact factor weighted publication on average increases the probability of selecting into the performance pay scheme by 4 or 5%. In contrast, for academics in the intermediate age bracket one extra impact factor weighted publication on average increases the probability of selecting into the performance pay scheme by only 0.6 or 0.9%, and this effect reduces to only 0.4% for academics with the maximum age-related pay level, in the highest age bracket. The negative and significant interaction term of age and productivity in the interaction-augmented Weibull proportional hazard models in Table 4.8 in appendix B confirms the finding that the effect of average productivity on the probability of selecting into performance pay decreases with age[29].

4.4 Conclusion

This chapter studies the effort and selection effect of performance pay in academia and provides empirical evidence that academics significantly increase effort in response to performance pay and that higher productivity academics are more likely to select into performance pay. In order to do so, I use the introduction of performance pay in German academia in 2002 as a natural experiment and employ a newly constructed data set encompassing information regarding research productivity and affiliations of the universe of German academics. Before the reform, academics were all paid according to an age-related pay scheme, in which the effectively flat wage increases with age. In contrast, in the performance pay scheme implemented after the reform, academics earn a basic wage that does not increase with age and is lower than the basic wage in the age-related pay scheme for most ages. On top of this basic wage however, academics can now earn bonuses that are partly distributed through on-the-job performance tournaments and partly through an, effectively, piece-rate scheme.

[29]Estimation of an equivalent Cox model yields equivalent results (results available upon request).

In order to estimate the effort effect, the chapter exploits the fact that academics who make tenure just before the reform fall under the old, age-related pay scheme, while academics who make tenure directly after the reform are paid according to the performance pay scheme. If the timing of the tenure decision is exogenous, the difference in the change in productivity from before to after the reform between the cohort making tenure just before the reform and the cohort making tenure directly after, can be interpreted as the causal effect of performance pay on academic effort. I estimate this differential change in a difference-in-differences framework and find an effort effect that amounts to a 35% increase in productivity. About one-third of this effort effect is caused by the piece-rate component of the performance pay scheme, with the remaining two-thirds instigated by the on-the-job performance bonuses. A placebo difference-in-differences estimation shows that the identifying assumption of an exogenous tenure decision is plausible, as there is no evidence of academics speeding up their tenure process. Furthermore, I find no evidence of pre-existing trends, which lends support to the identifying parallel-trends assumption of the difference-in-differences framework.

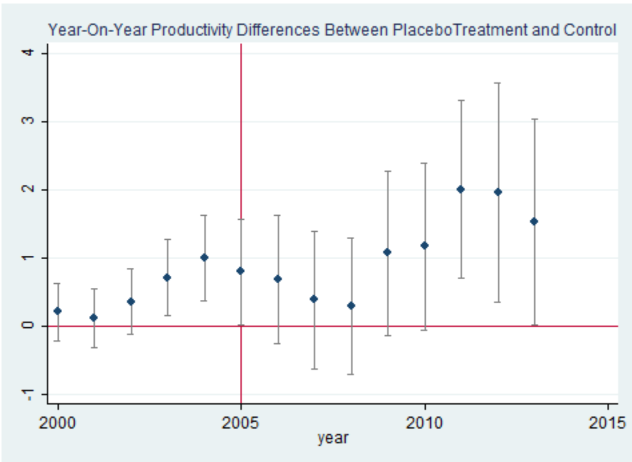
I estimate the selection effect though hazard rate and survival function analysis, where I use the fact that any tenured professor who changes affiliation or position after the reform automatically switches to the performance pay scheme. Because pay in the performance pay scheme only increases with performance, and no longer with age, I expect to find that more productive academics are more likely to switch to performance pay. This selection effect is borne out by the analysis indeed and, moreover, I find that this effect is stronger for younger academics. The latter finding aligns with the fact that the gap in basic wage between the age-related pay and the performance pay scheme is larger for older academics.

By studying the effort and selection effect of performance pay in academia, the chapter aims to contribute to and form a bridge between the literature on university governance and incentives in organizations. Given the economic importance of the academic sector, both in terms of direct economic value as well as for innovation and growth, it is crucial to understand the factors that determine the performance of universities. The literature on incentives in organisations provides ample evidence that incentives can significantly increase performance, and this chapter shows that incentives can improve academic performance too.

There are several steps than can be taken next. The current chapter focuses on research productivity of academics, but it would be very interesting and equally relevant to study the effect of performance pay on the educational performance of academics. Furthermore, it would be interesting to see if and when extrinsic motivation starts to crowd out academics' intrinsic motivation. Another important question is how performance pay affects the selection of candidate-academics. Other possible impacts of performance pay, such as on collaboration and network formation would make for exciting research avenues too.

Figure 4.4.1: Confidence Intervals of Year-on-Year Productivity Differences Between Treatment and Control Group

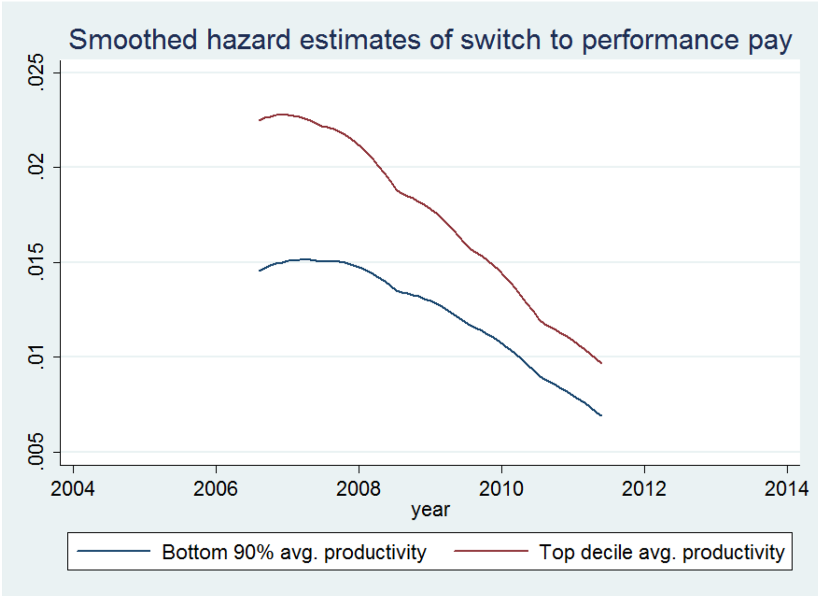
(a)



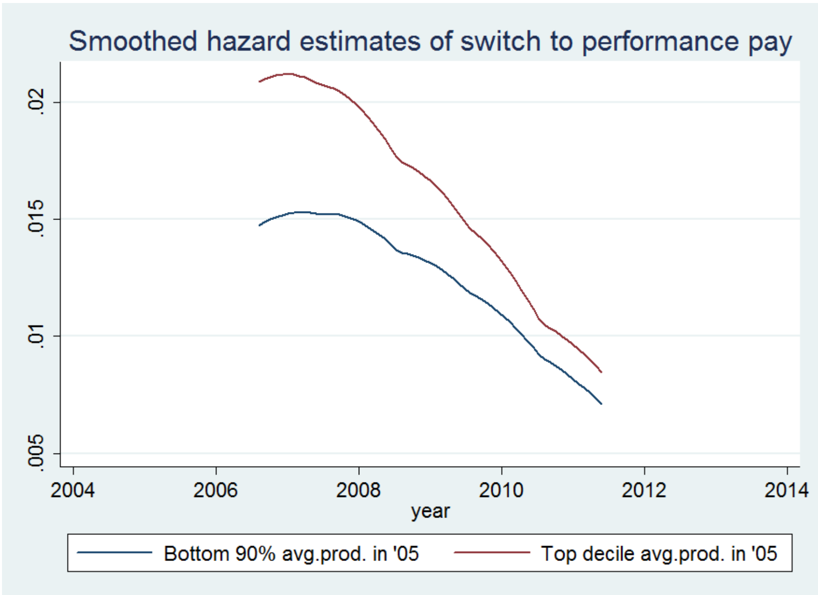
(b)

The figures depict the coefficient estimates and corresponding 90% confidence intervals of the interactions of a treatment dummy and year dummies in a regression of weighted publications in t on year dummies and year*treatment interactions, controlling for individual fixed effects. The dependent variable is the weighted number of publications of academic i in field f and year t, lagged by average publication lag in field f as reported in Björk and Solomon (2013). The unit of observation is academic i. In Figure a, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006, so that *Treatment* is 1 if an academic starts his first tenured affiliation in 2005 or 2006 and 0 if he starts his first tenured affiliation in 2003 or 2004 (the control group). In Figure b, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2001, 2002, 2003 or 2004, and *Treatment* is 1 if an academic starts his first tenured affiliation in 2003 or 2004 and 0 if he starts his first tenured affiliation in 2001 or 2002 (the control group).Standard errors are robust, clustered by individual academic.

Figure 4.4.2: Smoothed Hazard Estimates of Switch to Performance Pay



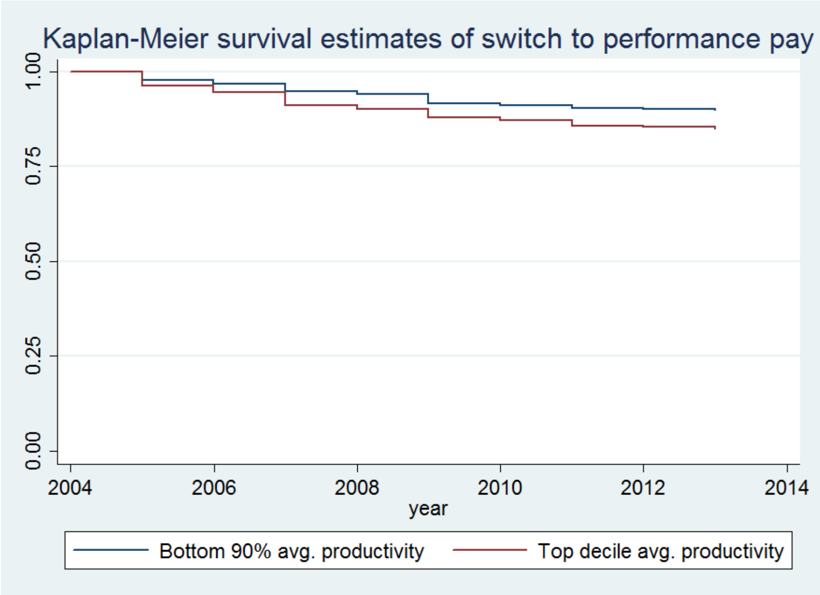
(a)



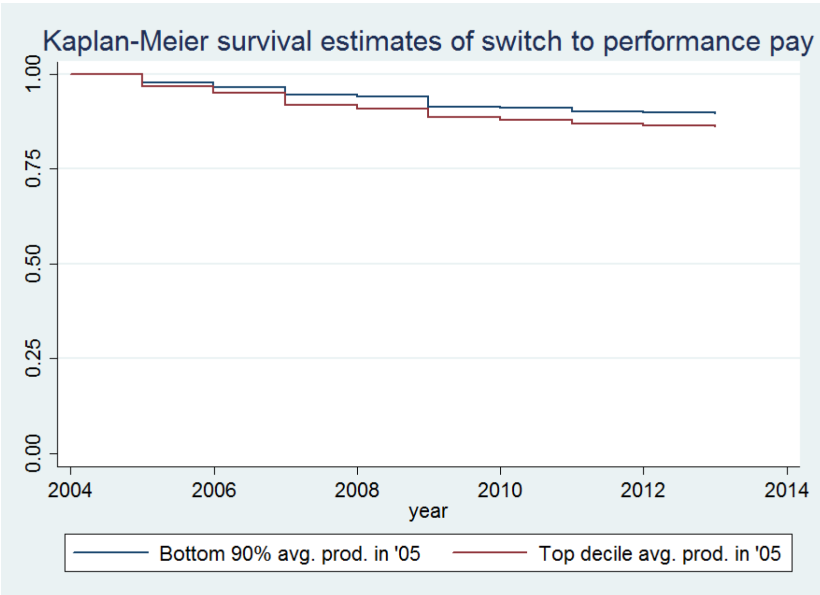
(b)

The above figures depict the Epanechnikov kernel-density estimates of the hazard function for switching to the performance pay scheme for academics in the top decile and bottom 90% of the average productivity distribution. Only switches to performance pay from age-related pay are considered, where first tenured affiliations after 2004 are not considered switches. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. In Figure a, an academic belongs to the top decile if his average productivity in year t falls in the top decile of the average productivity amongst all academics at public universities in year t. In Figure b, an academic belongs to the top decile if his average productivity in 2005 falls in the top decile of the average productivity amongst all academics at public universities in 2005.

Figure 4.4.3: Kaplan-Meier Survival Estimates of Switch to Performance Pay



(a)



(b)

The above figures depict the Kaplan-Meier estimates of the survival function for switching to the performance pay scheme for academics in the top decile and bottom 90% of the average productivity distribution. Only switches to performance pay from age-related pay are considered, where first tenured affiliations after 2004 are not considered switches. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. In Figure a, an academic belongs to the top decile if his average productivity in year t falls in the top decile of the average productivity amongst all academics at public universities in year t. In Figure b, an academic belongs to the top decile if his average productivity in 2005 falls in the top decile of the average productivity amongst all academics at public universities in 2005.

Table 4.1: Summary Statistics

	Mean	Std	Min	Max	Median
Panel A: Effort Effect Analysis					
Weighted Publications (All academics, 1999-13)	3.783	16.709	0	1298.662	0
Weighted Publications (At public uni post'05, 1999-13)	4.737	18.910	0	1298.662	0
Weighted Publications (Tenured at public uni post'05, 1999-13)	5.235	22.038	0	1298.662	0
Weighted Publications (1st Tenure at public uni 03-06, 1999-13)	7.333	26.594	0	705.255	0
Panel B: Selection Effect Regressions	**Total**	**Mean**	**Min**	**Median**	**Max**
Switches	3376	0.090	0	0	1
Time at risk	248107	6.604	1	9	9

Notes: The unit of observation is academic i. For the baseline effort effect estimations, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006. For the selection effect estimations, the sample is restricted to academics who made tenure before 2005.

Table 4.2: Unconditional Difference-in-Differences: Means of Weighted Number of Publications by First Tenure Cohort and Before/After Reform Period

	Treatment group (first tenured '05/'06)	**Control group (first tenured '03/'04)**	**Difference**
Before reform ('99-'04)	4.693	4.362	0.331
	(0.377)	(0.256)	(0.455)
After reform ('05-'13)	10.667	8.794	**1.873****
	(0.796)	(0.525)	**(0.953)**
Difference	**5.974*****	**4.432*****	**1.542****
	(0.540)	**(0.360)**	**(0.649)**

(a) Experiment of Interest

	Treatment group (first tenured '03/'04)	**Control group (first tenured '01/'02)**	**Difference**
Before reform ('99-'04)	4.362	4.196	0.166
	(0.256)	(0.236)	(0.348)
After reform ('05-'13)	8.794	7.931	0.863
	(0.525)	(0.509)	(0.731)
Difference	**4.432*****	**3.734*****	0.697
	(0.360)	**(0.362)**	(0.511)

(b) Placebo Experiment

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The table shows the means, differences in means and corresponding standard errors of the impact factor weighted number of publications of the treatment and control group before and after the performance pay reform. These means and differences in means are derived from a pooled OLS regression of the lagged, weighted number of publications of academics on a constant, a treatment dummy, a post dummy (that is 1 for the years after the reform (2005-2013), and 0 before) and a treatment*post interaction term, where the number of publications of academic i in field f and year t are lagged by the average publication lag in field f as reported in Björk and Solomon (2013). The before-reform period spans 6 years (1999-2004), the period after the reform spans 9 years (2005-2013). In Panel A, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006 with the treatment group comprising academics who first made tenure in 2005 or 2006, while academics who first made tenure in 2003 or 2004 make up the control group. In Panel B, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2001, 2002, 2003 or 2004, with the (placebo) treatment group comprising academics who first made tenure in 2003 or 2004, while academics who first made tenure in 2001 or 2002 make up the (placebo) control group. Standard errors are clustered by individual academic and reported in parentheses.

Table 4.3: Baseline Diff-in-Diff and Pre-Trend Test

	Baseline		Pre-trend Test	
Dep. Var.: Weighted Number of Publications	**1_a**	**1_b**	**2_a**	**2_b**
Linear Time Trend		**0.721*** **		**0.780*** **
		(0.056)		**(0.069)**
Post_t-3 (1 if year≥ 2002,0 otherwise)			**0.548*** **	**-1.012*** **
			(0.176)	**(0.225)**
Post_t-2 (1 if year≥ 2003,0 otherwise)			**0.822*** **	0.042
			(0.189)	(0.196)
Post_t-1 (1 if year≥ 2004,0 otherwise)			0.297	**-0.483** **
			(0.192)	**(0.205)**
Post (1 if year≥ 2005,0 otherwise)	**4.432*** **	**-0.975** **	**3.362*** **	-0.538
	(0.360)	**(0.411)**	**(0.343)**	(0.372)
Post_t-3 * Treat. (1st tenured '05/'06)			**0.543** **	**0.543** **
			(0.275)	**(0.275)**
Post_t-2 * Treatment			-0.186	-0.186
			(0.328)	(0.328)
Post_t-1 * Treatment			0.301	0.301
			(0.357)	(0.357)
Post * Treatment	**1.542** **	**1.542** **	**1.144* **	**1.144* **
	(0.649)	**(0.649)**	**(0.635)**	**(0.635)**
N	3717	3717	3717	3717

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is academic i. The sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006. The dependent variable is the impact-factor weighted number of publications of academic i in field f and year t, lagged by average publication lag in field f as reported in Björk and Solomon (2013). *Post* is 0 before 2005 and 1 thereafter, and *Treatment* is 1 if an academic starts his first tenured affiliation in 2005 or 2006 and 0 if he starts his first tenured affiliation in 2003 or 2004 (the control group). $Post\_t-3$ is 0 before 2002 and 1 thereafter, $Post\_t-2$ is 0 before 2003 and 1 thereafter and $Post\_t-1$ is 0 before 2004 and 1 thereafter. All regressions contain individual fixed effects. Standard errors are robust, clustered by individual academic.

Table 4.4: Placebo Diff-in-Diff and Pre-Trend Test

	Trend Break DiD		Pre-trend Test	
Dep. Var.: Weighted Number of Publications	**1_a**	**1_b**	**2_a**	**2_b**
Linear Time Trend		**0.539*** **		**0.594*** **
		(0.041)		**(0.050)**
Post_t-3 (1 if year≥ 2002,0 otherwise)			**0.299***	**-0.889*** **
			(0.165)	**(0.201)**
Post_t-2 (1 if year≥ 2003,0 otherwise)			**0.468** **	-0.126
			(0.206)	(0.206)
Post_t-1 (1 if year≥ 2004,0 otherwise)			0.017	**-0.577*** **
			(0.186)	**(0.198)**
Post (1 if year≥ 2005,0 otherwise)	**3.734*** **	-0.305	**3.259*** **	0.290
	(0.362)	(0.389)	**(0.335)**	(0.344)
Post_t-3 * Placebo-Treat. (1st tenured '03/'04)			0.249	0.249
			(0.241)	(0.241)
Post_t-2 * Placebo-Treatment			0.354	0.354
			(0.279)	(0.279)
Post_t-1 * Placebo-Treatment			0.281	0.281
			(0.267)	(0.267)
Post * Placebo-Treatment	0.697	0.697	0.103	0.103
	(0.511)	(0.511)	(0.479)	(0.479)
N	4270	4270	4270	4270

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is academic i. The sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006. The dependent variable is the weighted number of publications of academic i in field f and year t, lagged by average publication lag in field f as reported in Björk and Solomon (2013). *Post* is 0 before 2005 and 1 thereafter, and *Treatment* is 1 if an academic starts his first tenured affiliation in 2003 or 2004 and 0 if he starts his first tenured affiliation in 2001 or 2002 (the control group). $Post\_t-3$ is 0 before 2002 and 1 thereafter, $Post\_t-2$ is 0 before 2003 and 1 thereafter and $Post\_t-1$ is 0 before 2004 and 1 thereafter. All regressions contain individual fixed effects. Standard errors are robust, clustered by individual academic.

Table 4.5: Proportional Hazard Model

	1_a	**1_b**	**2_a**	**2_b**
Average Productivity	**0.003*****	**0.005*****	**0.003*****	**0.005*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
Age		**-0.092*****		**-0.092*****
		(0.003)		**(0.003)**
Constant	**-4.304*****	-0.015	**-4.304*****	-0.015
	(0.027)	(0.160)	**(0.027)**	(0.160)
$ln(\rho)$	-0.006	**0.145*****	-0.003	**0.148*****
	(0.010)	**(0.011)**	(0.010)	**(0.011)**
N	37571	37562	37571	37562

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is academic i. The duration under consideration is the time until switch to performance pay, and academics are considered "at risk" of switching after they have made tenure. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. In columns 1a and 1b, the average productivity in year t is used, while in columns 2a and 2b the average productivity in 2005 is used as covariate. The age variable is equal to an author's self-reported age if known, and equal to a synthetic age otherwise. The synthetic age is calculated using the average age at habilitation and start of academic career. Standard errors are robust, clustered by individual academic.

Table 4.6: Heterogeneous Hazard Rates by Age - Split Sample

	Age $<$ 37		**37 $\leq$ Age $<$ 49**		**Age $\geq$ 49**	
	1_a	**1_b**	**2_a**	**2_b**	**3_a**	**3_b**
Average Productivity	**0.040*** **	**0.053*** **	**0.006*** **	**0.009*** **	**0.004*** **	**0.004*** **
	(0.006)	**(0.009)**	**(0.001)**	**(0.002)**	**(0.001)**	**(0.001)**
Constant	**-3.414*** **	**-3.466*** **	**-4.220*** **	**-4.232*** **	**-4.701*** **	**-4.700*** **
	(0.153)	**(0.158)**	**(0.039)**	**(0.040)**	**(0.045)**	**(0.045)**
$ln\,(\rho)$	-0.075	-0.043	**0.212*** **	**0.216*** **	-0.012	-0.008
	(0.046)	(0.046)	**(0.015)**	**(0.015)**	(0.017)	(0.017)
N	463	463	19654	19654	30695	30695

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. All models estimated are Weibull models. The unit of observation is academic i. The duration under consideration is the time until switch to performance pay, and academics are considered "at risk" of switching after they have made tenure. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. In the a columns the average productivity in year t is used, while in the b columns the average productivity in 2005 is used as covariate. The sample is restricted to academics at risk of switching while younger than 37 in columns 1a and 1b, between the ages of 37 and 48 in column 2a and 2b, and 49 or older in columns 3a and 3b. Standard errors are robust, clustered by individual academic.

Appendix A

Proof of Proposition 1:

Comparing the FOC under performance pay 4.2.7 with that under age-related pay 4.2.5, and using that $C" > 0$ and $f_1 > 0$, it follows that equilibrium effort under performance pay is greater than under age-related pay if the marginal pay-off from effort is greater under performance pay than under age-related pay. That is, if:

$$\left[r^p + bg_{\varepsilon_j - \varepsilon_i}\left(\triangle f_{ij}\left(e^{p*};\theta\right)\right)\right] > r^a \tag{4.4.1}$$

This condition is met because, by assumption, $r^p > r^a$, $b > 0$ and $g\left(\varepsilon_j - \varepsilon_i\right) > 0$ everywhere on its support. Because I solve for subgame perfect equilibria, and because the optimisation problem in the effort stage is the same in each period[30], the equilibrium effort levels defined in 4.2.5 and 4.2.7 also give first period equilibrium effort. QED

Proof of Proposition 2

I first derive the sorting behaviour of academics in period 2, followed by that in period 1 to prove the statements regarding the heterogeneity of the selection effect across ages. The statement regarding the overall selection effect immediately follows from this.

In period 2, an academic's expected life-time equilibrium utility when working under the age-related pay scheme is given by:

$$E\left[U^a_{i,2}\left(e^{a*}_{i,2};\theta_i\right)\right] = w^a_2 + r^a f\left(e^{a*}_{i,2};\theta_i\right) - C\left(e^{a*}_{i,2}\right) \tag{4.4.2}$$

while his expected life-time utility when working under the performance pay scheme is given by:

$$E\left[U^p_{i,2}\left(e^{p*}_{i,2};\theta_i\right)\right] = w^p_2 + r^p f\left(e^{p*}_{i,2};\theta_i\right) + bG_{\varepsilon_j - \varepsilon_i}\left(\triangle f_{ij}\left(e^{p*}_2;\theta\right)\right) - C\left(e^{p*}_{i,2}\right) \tag{4.4.3}$$

By 4.4.1, the marginal pay-off from effort is larger under performance pay than under age-related pay at any given effort level. Furthermore, from proposition 1 we have that $e^{p*}_{i,2} > e^{a*}_{i,2}$. It then follows from the second order condition 4.2.8 that the portion of utility that varies with effort is greater under performance pay than under age-related pay. That

[30] Only the age-related basic wage is different in the first period, but because this does not depend on effort, the optimisation problem with respect to effort is the same.

is

$$r^p f\left(e_{i,2}^{p*};\theta_i\right)+bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_2^{p*};\theta\right)\right)-C\left(e_{i,2}^{p*}\right)>r^a f\left(e_{i,2}^{a*};\theta_i\right)-C\left(e_{i,2}^{a*}\right) \tag{4.4.4}$$

However, in period 2 the basic wage under age-related pay is larger than that under performance pay ($w_2^a > w_2^p$), so that the total expected utility under performance pay might be larger or smaller than under age-related pay.

By the envelope theorem and the assumptions that $f_1 > 0$, $r^p > r^a$, $b > 0$ and $g\left(\varepsilon_j-\varepsilon_i\right) > 0$ everywhere on its support, we have that equilibrium utility is larger for higher ability academics under both pay schemes:

$$\frac{\partial E\left[U_{i,2}^p\left(e_{i,2}^{p*};\theta_i\right)\right]}{\partial\theta_i}=\left[r^p+bg_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_2^{p*};\theta\right)\right)\right]\frac{\partial f\left(e_{i,2}^{p*};\theta_i\right)}{\partial\theta_i}>0$$

and

$$\frac{\partial E\left[U_{i,2}^a\left(e_{i,2}^{a*};\theta_i\right)\right]}{\partial\theta_i}=r^a\frac{\partial f\left(e_{i,2}^{a*};\theta_i\right)}{\partial\theta_i}>0$$

Furthermore, it follows from 4.4.1 that:

$$\frac{\partial E\left[U_{i,2}^p\left(e_{i,2}^{p*};\theta_i\right)\right]}{\partial\theta_i}>\frac{\partial E\left[U_{i,2}^a\left(e_{i,2}^{a*};\theta_i\right)\right]}{\partial\theta_i} \tag{4.4.5}$$

Define θ_2^0 to be the ability level at which $E\left[U_{i,2}^p\left(e_{i,2}^{p*};\theta_2^0\right)\right]=E\left[U_{i,2}^a\left(e_{i,2}^{a*};\theta_2^0\right)\right]$, so that academics with ability θ_2^0 are indifferent between the age-related pay scheme and the performance pay scheme. By 4.4.5, it must be that academics with ability level $\theta > \theta_2^0$ prefer performance pay to age-related pay, and hence self-select into the performance pay scheme. Then if $\theta_2^0 > \underline{\theta}$, only relatively high ability academics select into performance pay. The condition $\theta_2^0 > \underline{\theta}$ is met if the difference in basic wage between the age-related pay scheme and the performance pay scheme is sufficiently large, in particular, if the following holds:

$$w_2^a-w_2^p>r^p f\left(e_{i,2}^{p*};\theta_i\right)+bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_2^{p*};\theta\right)\right)-C\left(e_{i,2}^{p*}\right)-\left[r^a f\left(e_{i,2}^{a*};\theta_i\right)-C\left(e_{i,2}^{a*}\right)\right] \tag{4.4.6}$$

In period 1, an academic's expected life-time equilibrium utility when working under the age-related pay scheme in both periods is given by:

$$E\left[U_{i,1}^{a}\left(e_{i,1}^{a*};\theta_i\right)\right] = w_1^a + r^a f\left(e_{i,1}^{a*};\theta_i\right) - C\left(e_{i,1}^{a*}\right) + \delta\left[w_2^a + r^a f\left(e_{i,2}^{a*};\theta_i\right) - C\left(e_{i,2}^{a*}\right)\right] \quad (4.4.7)$$

while his expected lifetime utility when working under the performance pay scheme in both periods is given by:

$$E\left[U_{i,1}^{p}\left(e_{i,1}^{p*};\theta_i\right)\right] = \bar{w}_1^p + r^p f\left(e_{i,1}^{p*};\theta_i\right) + bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_1^{p*};\theta\right)\right) - C\left(e_{i,1}^{p*}\right) +$$

$$\delta\left[w_2^p + r^p f\left(e_{i,2}^{p*};\theta_i\right) + bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_2^{p*};\theta\right)\right) - C\left(e_{i,2}^{p*}\right)\right] \quad (4.4.8)$$

In a pure strategy subgame perfect Nash equilibrium, $e_{i,1}^{a*} = e_{i,2}^{a*}$and $e_{i,1}^{p*} = e_{i,2}^{p*}$. From 4.4.4, we therefore have that the portion of utility that varies with effort is greater under performance pay than under age-related pay in both period 1 and period 2. On the other hand, the difference between the basic wage in the age-related pay system and the performance pay system is larger positive in period 2 than in period 1. Period 2 utility is however discounted by δ. Denoting the difference in the portion of utility that varies with effort between the age-related and performance pay scheme by $\triangle u_2\left(e_2^*;\theta_i\right) \equiv r^p f\left(e_{i,2}^{p*};\theta_i\right) + bG_{\varepsilon_j-\varepsilon_i}\left(\triangle f_{ij}\left(e_2^{p*};\theta\right)\right) - C\left(e_{i,2}^{p*}\right) - \left[r^a f\left(e_{i,2}^{a*};\theta_i\right) - C\left(e_{i,2}^{a*}\right)\right] = \triangle u_1\left(e_1^*;\theta_i\right)$, we have that the difference in expected utility between working under the performance pay scheme and working under the age-related pay scheme in both periods is:

$$E\left[U_{i,1}^{p}\left(e_{i,1}^{p*};\theta_i\right)\right] - E\left[U_{i,1}^{a}\left(e_{i,1}^{a*};\theta_i\right)\right] = \bar{w}_1^p - w_1^a + \triangle u_1\left(e_1^*;\theta_i\right) + \delta\left[w_2^p - w_2^a + \triangle u_2\left(e_2^*;\theta_i\right)\right] \quad (4.4.9)$$

An academic prefers to switch to the performance pay scheme at the beginning of period 1 if 4.4.9 is positive. Let θ_1^0 be the ability level at which $E\left[U_{i,1}^{p}\left(e_{i,1}^{p*};\theta_1^0\right)\right] = E\left[U_{i,1}^{a}\left(e_{i,1}^{a*};\theta_1^0\right)\right]$. Academics with ability θ_1^0 are indifferent between switching from the age-related pay scheme to the performance pay scheme in period 1. By definition of θ_2^0 we have that $w_2^a - w_2^p + \triangle u_2\left(e_2^*;\theta_2^0\right) = 0$. Because $\triangle u_1\left(e_1^*;\theta_i\right) = \triangle u_2\left(e_2^*;\theta_i\right)$ and $\bar{w}_1^p - w_1^a > w_2^p - w_2^a$, and using 4.4.5 it then follows that $\theta_1^0 < \theta_2^0$. Therefore, there are academics who prefer to switch to performance pay in period 1, whose ability is not large enough for them to prefer switching in period 2. If $\theta_1^0 > \underline{\theta}$, it is still true that only relatively higher ability academics switch to performance pay in period 1, but amongst these young academics there are some with ability $\theta_1^0 < \theta < \theta_2^0$, who do prefer to switch when they are young (in period 1), but who would not switch when they are older (in period 2). Incidentally, this also implies that academics who do not switch to performance pay when they are young (in period 1), will not switch in period 2 either, since they must have ability $\theta < \theta_1^0 < \theta_2^0$. Finally because academics cannot switch back to age-related pay once they are in the performance pay scheme, we do not have to analyse the scenario in which academics select into the performance pay scheme in period 1, then switch back to age-related pay in period 2, when the basic wage difference between age related pay and performance pay is relatively larger. QED

Appendix B

Table 4.7: Year-on-Year Differences in Academic Productivity

Interaction-Year	Experiment of Interest	Placebo Experiment
2000	-0.246	0.200
	(0.271)	(0.253)
2001	0.177	0.116
	(0.304)	(0.262)
2002	0.520	0.354
	(0.329)	(0.293)
2003	0.334	**0.708****
	(0.402)	**(0.343)**
2004	0.635	**0.989*****
	(0.462)	**(0.380)**
2005	**0.967***	**0.792***
	(0.510)	**(0.469)**
2006	**1.407****	0.672
	(0.692)	(0.572)
2007	**1.352****	0.378
	(0.687)	(0.613)
2008	**1.429****	0.282
	(0.664)	(0.607)
2009	**1.634***	1.071
	(0.935)	(0.734)
2010	1.285	1.166
	(0.937)	(0.747)
2011	1.578	**2.001****
	(1.016)	**(0.791)**
2012	**2.680****	**1.951****
	(1.365)	**(0.980)**
2013	**3.680*****	**1.514***
	(1.322)	**(0.919)**
N	3717	4270

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is academic i. The table reports the coefficients of the interactions of a treatment dummy and year dummies in a regression of weighted publications in t on year dummies and year*treatment interactions, controlling for individual fixed effects. The dependent variable is the weighted number of publications of academic i in field f and year t, lagged by average publication lag in field f as reported in Björk and Solomon (2013). In column 1, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2003, 2004, 2005 or 2006, so that *Treatment* is 1 if an academic starts his first tenured affiliation in 2005 or 2006 and 0 if he starts his first tenured affiliation in 2003 or 2004 (the control group). In column 2, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2005, 2006, 2007 or 2008, and *Treatment* is 1 if an academic starts his first tenured affiliation in 2007 or 2008 and 0 if he starts his first tenured affiliation in 2005 or 2006 (the control group). In column 3, the sample is restricted to academics who started their first tenured affiliation at a German public university in 2001, 2002, 2003 or 2004, and *Treatment* is 1 if an academic starts his first tenured affiliation in 2003 or 2004 and 0 if he starts his first tenured affiliation in 2001 or 2002 (the control group). Standard errors are robust, clustered by individual academic.

Table 4.8: Heterogeneous Hazard Rates by Age - Interactions

	Weibull Model	
	1a	**1b**
Average Productivity	**0.025*** **	**0.034*** **
	(0.007)	**(0.010)**
Age	**-0.090*** **	**-0.090*** **
	(0.004)	**(0.004)**
Average Productivity*Age	**-0.0004*** **	**-0.001*** **
	(0.000)	**(0.000)**
Constant	-0.133	-0.129
	(0.163)	(0.163)
ln(p)	**0.144*** **	**0.149*** **
	(0.011)	**(0.011)**
N	37562	37562

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is academic i. The duration under consideration is the time until switch to performance pay, and academics are considered “at risk” of switching after they have made tenure. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. In the a columns the average productivity in year t is used, while in the b columns the average productivity in 2005 is used as covariate. Standard errors are robust, clustered by individual academic.

5 Lone Stars or Constellations? The Impact of Performance Pay on Matching Assortativeness in Academia

5.1 Introduction

Performance related pay is widespread and becoming increasingly more prevalent. Using PSID data, Lemieux et al. (2009) show that the incidence of performance pay jobs among US salaried workers was less than 45% in the late 1970s and increased to almost 60% at the end of the 1990s. While the effects of performance pay on on-the-job performance and sorting are widely studied[1], the effect performance pay may have on team composition in general, and the degree of matching assortativeness in particular, has not been studied extensively. The level of matching assortativeness may however greatly affect total productivity in sectors and countries. In his seminal "O-Ring Theory" paper, Kremer (1993) for instance shows that if production is complementary in worker skill, workers will be matched positive assortatively by skill; output and wages increase sharply in skill; and incomes differ greatly between countries. In this chapter I study the effect of performance pay on matching assortativeness and provide empirical evidence that performance pay increases positive assortative matching by productivity if there are positive productivity spillovers, using the introduction of performance pay in German academia as a natural experiment.

I present a simple matching model that makes precise how performance related pay affects matching assortativeness. I model the academic job market as a stochastic hedonic coalition formation problem in which an academic's utility from a coalition depends on a systematic and an idiosyncratic component. The systematic component depends on the

[1] See Lazear and Oyer (2012), Bloom and Reenen (2011), Oyer and Schaefer (2011), Lazear and Shaw (2007) for an overview of the literature.

productivity of the academics in a coalition and increases in own and partner's productivity to represent spillover effects. As in Choo and Siow (2006) and Siow (2009)'s stochastic Becker (1973) model, the idiosyncratic component captures the deviation of an academic's utility from the systematic component of the utility from a given coalition and reconciles observing matchings that are not perfect positive assortative matchings by productivity in the academic job market with the theory. In the model, the idiosyncratic component captures a worker's personal preferences regarding colleagues and workplace, and it causes matching to become less positively assortative. This reduces total output if the academic output function exhibits increasing differences. The introduction of performance pay increases the utility from academic output and thus increases positive assortative matching by productivity again. Moreover, this increase is larger if complementarities are stronger. I test this hypothesis empirically in a difference-in-differences framework, using the introduction of performance pay in German academia as a natural experiment and the strength of complementarities in academic fields as a measure of treatment intensity.

Any appointment after the implementation of the reform necessarily falls under the new performance pay scheme (Detmer and Preissler 2004). The fact that this is a nationwide reform that affects all agents in an entire sector means that agents cannot simply opt into or out of either one of the pay schemes by moving to a different employer[2]. Academics who already had a tenured position can of course avoid the new pay scheme by staying put, and any (aspiring) academics can avoid the performance pay scheme by leaving academia, but any academic affiliation decisions made after the reform are made in the face of the new performance pay scheme and therefore under the influence of incentives to match more or less assortatively that come with performance pay.

In order to study the effect of performance pay on matching assortativeness empirically, I employ the department level data set that I derived from the individual level data set comprising the affiliations, productivity and related information of the universe of academics in Germany which I constructed for this research project. I use this data set to study any changes in departmental composition from before to after the reform. The combination of a data set that encompasses an entire sector in a country and a reform that introduces performance pay throughout the same sector in the country allows for estimation of the effect of performance pay on matching assortativeness[3].

[2]As long as they remain employed in the academic sector and do not move abroad.

[3]To my knowledge, this is the first study into the German academic pay reform using a data set that

The empirical analysis comprises of two stages. I first assess whether positive assortative matching increases post-reform and subsequently test whether it is in fact performance pay that increases positive assortativeness matching. I do so by studying the two channels through which departmental composition can change: hiring and firing. As for hiring, I make a distinction between junior and senior hiring, thus providing a further insight in the anatomy of compositional changes of departments.

I find that positive assortative matching by productivity increases post-reform: the difference in average productivity of new hires between high quality and low quality departments more than doubles, while the difference in average productivity of leavers between high and low quality departments decreases by more than half. I then test whether performance pay increases positive assortative matching if there are complementarities in worker skill by comparing the difference in changes in positive assortativeness in fields with weaker and stronger complementarities in a difference-in-differences framework, using the strength of complementarity in a field as a measure of treatment intensity. If performance pay increases positive assortative matching, the increase should be larger in fields in which complementarities are stronger. I find that the difference in average productivity of new hires between high quality and low quality departments is more than three times larger post-reform in fields in which complementarities are stronger, so the increase in positive assortativeness after the reform is considerably larger in high complementarity fields. This is consistent with performance pay increasing positive assortativeness as driven by complementarities in worker productivity.

I control for alternative explanations, such as pre-existing trends and differential hiring budgets and show that the results are robust. Because controlling for a department's hiring budget with a contemporaneous hiring budget variable comes with the risk of omitted variable bias, I construct a proxy for the hiring budget that is historically determined and plausibly exogenous. Since a German university's personnel budget varies relatively little from one year to the next (Jongbloed 2009) and academics are mandated to retire at 65 (Mohr 2007, *Bundesbeamtengesetz* 1985), and because academics who are about to retire earn the highest salary under the old, age-related pay scheme, if a lot of academics retire in a given year, a larger share of the personnel budget is available for hiring. A university's hiring budget thus varies with the number of academics who retire in a year and this

encompasses the universe of academics in Germany.

variation is historically determined and plausibly exogenous.

To my knowledge this study is one of the first to investigate the effect of performance pay on the level of assortative matching in the presence of complementarities in worker skill. Bandiera et al. (2013) study the effect of team-based incentives on matching assortativeness in teams within a firm employing manual labourers. This chapter studies the effect of individual pay for performance incentives on the assortativeness of matching into firms (universities) in an entire sector in a country where the workers involved are knowledge workers and the production function is thought to exhibit complementarities in worker skill.

Studying the effect of performance pay on matching assortativeness in the context of academia is interesting and relevant for a number of reasons. Various papers show that there can be significant spillover effects in academia, not only for co-authors (Borjas and Doran 2014, Oettl 2012, Waldinger 2012, Azoulay et al. 2010) and PhD's (Waldinger 2012, 2010), but also for co-located colleagues who work on related subjects (Agrawal et al. 2014)[4]. Furthermore, by studying the effect of performance pay on assortativeness in academia, this chapter aims to add to the literature on the organisation of knowledge creation (Grigoriou and Rothaermel 2014, Phelps et al. 2012, Jones 2009, Wuchty et al. 2007, Singh 2005, Audretsch and Feldman 2004, 1996, Jaffe et al. 1993) and the debate on university governance (cf. Haeck and Verboven 2012, Aghion et al. 2010, Belenzon and Schankerman 2009, Lach and Schankerman 2008, 2004) in particular. Given that human capital organisation is of primary importance in knowledge creation, academia seems an important place to study the organisation of labour and management practices. Human capital and knowledge creation, and in particular human capital spillovers have, in turn, been understood to play a central role in economic growth since the models of Romer (1990) and (1986) and Lucas (1990) and (1988). Finally, studying matching in an entire sector in a country, as I do here, is relevant not just for the effect matching assortativeness may have on total output, but also the different distribution of output and production factors that a change in matching assortativeness implies. This distribution may affect welfare directly too and academia is a particularly relevant example of a sector in which both the total output and the distribution of production factors and output matter for total welfare. A

[4] Kim et al. (2009) however show that the externality of productive academic colleagues has diminished over the last three decades of the 21st century and even disappeared in the 1990s in the fields of economics and finance.

greater total scientific output may boost technological progress, so to the extent that there are positive productivity spillovers[5] in academia, this calls for a concentration of the most productive academics. On the other hand, we may also care about, for instance, providing good scientific education to many people, all over a country. This requires a more even distribution of high quality academics across a country. Hence if performance-related pay affects the matching assortativeness of academics (or productive agents in general) and hence their distribution, this may affect welfare in more ways than the direct effect of (potentially) increasing productivity. It is therefore especially relevant to study the effect of performance related pay on matching assortativeness in sectors, such as academia, in which the distribution of the production factors (and output) is an important determinant of welfare as well.

The rest of the chapter is organised as follows; I present the theoretical framework in the next section, describe the data in section 3 and discuss the empirical results in section 4. Section 5 concludes.

5.2 Model

In this section I present a simple one-sided matching model of team (in the context of the chapter: department) formation that makes precise the effects of performance related pay on matching assortativeness. In this model, whenever two academics are matched, they form a department and are considered active in academia. An academic who remains unmatched is considered to leave academia. This academic job market is akin to the roommate market introduced by Gale and Shapley (1962), in which a set of students is partitioned into pairs of roommates and singletons. As remarked by Klaus et al. (2010), these markets are instances of hedonic coalition formation problems where coalitions are restricted to comprise at most two agents.

Most roommate market and more general hedonic coalition formation problems are described by deterministic models (see e.g. Hajduková 2006). I will however model the academic job market as a stochastic hedonic coalition formation problem in which an academic's utility from a coalition depends on a systematic and an idiosyncratic component. The systematic component depends on the productivity of the academics in a coalition. As in Choo and Siow (2006) and Siow (2009)'s stochastic Becker (1973) model, the id-

[5]Specifically, if the output function is supermodular.

iosyncratic component captures the deviation of an academic's utility from the systematic component of the utility from a given coalition. This idiosyncratic component allows for observing matchings that are not perfect positive assortative matchings by productivity in the academic job market.[6] Moreover, it allows for increases in the level of positive assortative matching by productivity post-reform if the introduction of performance related pay increases the utility from the systematic component.

This model is similar to that presented in Bandiera et al. (2013) in that it models team-formation as a one-sided matching model without transferable utility. In that paper however, the utility of agents depends positively on their partner's productivity due to team-based incentive schemes and because of friendship ties, whereas in this chapter it depends positively on partner's productivity regardless of their relationship and due to positive spillovers in the production function. The model in Bandiera et al. (2013) also includes a component that does not depend on agents' productivity, but in contrast to the idiosyncratic component in this chapter, theirs is deterministic and depends on friendship ties. The idiosyncratic component considered here is stochastic and more general in that it may differ within pairs of agents so as to allow for any kind of unobservable preferences regarding potential matching partners' identity[7].

5.2.1 Model Set-Up - Baseline

Let there be m academics indexed $1, ..., m$. Academic i has productivity type θ_i, which is a random and independent draw from $\theta \sim U\left[\underline{\theta}, \overline{\theta}\right], \underline{\theta} = 1$. I assume that academics' productivity is common knowledge, because in the time period considered for the empirical analysis important indicators of an academic's productivity, such as educational and professional background, publication record and other academic achievements are readily available online for most academics. Let a faculty consist of two academics, so that when i and j are matched, they form a faculty and work in academia. If i remains unmatched, he leaves academia. The utility of academic i is given by:

[6] A definition of a positive assortative matching in the context of the one-sided matching model considered here is given shortly.

[7] This is because affiliation decisions are likely influenced by more than just friendship ties, such as the geographical location of a department, and preferences may be asymmetric.

$$u_{i|j} = \begin{cases} \alpha f\left(\theta_i \mid \theta_j\right) + w_{i|j} & \text{if } i \text{ is matched with academic } j \\ \underline{u}_i & \text{if } i \text{ is unmatched} \end{cases} \tag{5.2.1}$$

where $f\left(\theta_i \mid \theta_j\right)$ is the productivity of i when he is matched with j, $\alpha > 0$ is a weighing constant that captures academics' valuation of this productivity (for instance capturing intrinsic motivation) and $w_{i|j}$ is the wage i is paid when matched with j. Pre-reform, every academic receives the same, flat wage at a given age irrespective of the identity or productivity of his matching partner, so $w_{i|j} = \bar{w}_i$. For now, let every agent also have the same outside option, $\underline{u}_i = \underline{u}, \forall i$.

I assume that the production function $f\left(\theta_i \mid \theta_j\right)$ increases in own and matching partner's productivity. This increase in partner's productivity represents spillovers between academics and they are larger, the larger is the increase in output with respect to partner's productivity[8]. Specifically, let $f_1 = \frac{\partial f(\theta_i|\theta_j)}{\partial \theta_i} > 0$ and $f_2 = \frac{\partial f(\theta_i|\theta_j)}{\partial \theta_j} > 0$. The model is intentionally kept sufficiently general to allow for spillovers that are not restricted to co-authorships, but simply occur between academics who are working in the same department.

Define the matching (or assignment) function μ to be a 1-to-1 correspondence from $\{1, ..., m\} \rightarrow \{1, ..., m\}$ of order 2, so that:

$$\mu\left(i\right) = j \text{ if } i \text{ is matched with academic } j \tag{5.2.2}$$

Matchings are *symmetric*: if $\mu\left(i\right) = j$ then also $\mu\left(j\right) = i$. For ease of notation, I will use the shorthand μ_{ij} to denote a matching of i with j, where $\mu_{ij} = 1$ if $\mu\left(i\right) = j$. A matching μ is *feasible* if every academic is matched to one and only one academic (possibly himself), so that $\sum_j \mu_{ij} = 1$. A matching μ is *individually rational* if no academic prefers working outside of academia to being in a department with his matching partner in μ, that is if $u_{i|\mu(i)} = \alpha f\left(\theta_i \mid \theta_{\mu(i)}\right) + w_i \geq \underline{u}, \forall i \in \{1, .., m\}$. A matching μ cannot be improved upon if there are no two academics i, j such that

$$\alpha f\left(\theta_i \mid \theta_j\right) \geq \alpha f\left(\theta_i \mid \theta_{\mu(i)}\right)$$

[8] Formally, academic output function $f\left(. \mid .\right)$ exhibits stronger complementarities than $\tilde{f}\left(. \mid .\right)$ if $f\left(\theta_i \mid \theta_j\right)$- $f\left(\theta_i \mid \theta_k\right) > \tilde{f}\left(\theta_i \mid \theta_j\right) - \tilde{f}\left(\theta_i \mid \theta_k\right)$ for $\theta_j > \theta_k$

and

$$\alpha f(\theta_j \mid \theta_i) \geq \alpha f\left(\theta_j \mid \theta_{\mu(j)}\right)$$

with at least one inequality strict (i.e. no coalition can improve upon the allocation; no blocking pairs exist).

Following Gale and Shapley (1962), I define a matching μ to be *stable* if no academic or pair of academics wants to deviate from (or "block") it. More precisely, and following Chiappori et al. (2014), a matching of academics μ is *stable* if it is *feasible* and cannot be improved upon. A matching μ is *optimal* if it maximises total surplus, such that $\sum_i u_{i|\mu(i)} \geq \sum_i u_{i|\mu'(i)}, \forall \mu'(i)$.

Proposition 1 (Baseline matching): *The baseline model has a unique stable matching, which matches the highest productivity type with the next highest productivity type academic, and so on. This stable matching is optimal. If $f(\theta_i \mid \theta_j)$ exhibits increasing differences this is the unique optimal matching*[9]*.* (Proof in Appendix A)

I will refer to this matching $\tilde{\mu}$ as a *maximal positive assortative matching* of the academic job market, where the definition of such a matching coincides with that of the matching $\tilde{\mu}$. That is, I will refer to a matching μ as a *maximal positive assortative matching* of the academic job market if it matches the most productive academic with the second highest productivity academic, the third with the fourth, and so on, for all academics and their productivity adjacent match partners whose utility from such a match is at least as large as the utility from the outside option. Put differently, a matching μ is a *maximal positive assortative matching* if the (average) difference in productivity rank between two matched academics who are active in academia is 1.

Gale and Shapley (1962) showed that the core of two-sided one-to-one games (the marriage market) is non-empty, and that their proposed "deferred-acceptance" algorithm yields not only a stable, but an optimal assignment of agents. They also show that a stable matching may not exist in the one-sided roommate matching problem without transferable utility. From Shapley and Shubik (1971) we know that for bipartite matchings with transferable utility the set of stable allocations also coincides with the core, so that a stable matching μ is an optimal matching. This result does not necessarily carry over to one-sided matching

[9]The function $f(\theta_i \mid \theta_j)$ exhibits *increasing differences* if, for any $\theta_i' > \theta_i$ and $\theta_j' > \theta_j$ we have that $f\left(\theta_i' \mid \theta_j'\right) - f\left(\theta_i \mid \theta_j'\right) \geq f\left(\theta_i' \mid \theta_j\right) - f(\theta_i \mid \theta_j)$

problems with transferable utility, in which surplus-maximising matchings may not be stable (Chiappori et al. 2014). Proposition 1 shows that the baseline model presented here of one-sided team formation does have a unique stable matching and, moreover, that this stable matching is optimal too. Furthermore, if the matching production function $f(\theta_i \mid \theta_j)$ is supermodular, the unique stable matching is also the unique optimal matching. Hence if the academic output function exhibits increasing differences, the maximal positive assortative matching $\tilde{\mu}$ uniquely maximises total academic output. A distortion of the model that renders $\tilde{\mu}$ unstable therefore reduces total academic output[10]. In the next section I introduce such a distortion. On the other hand, an innovation that counters the effect of the distortion might increase the probability that $\tilde{\mu}$ is stable again and total academic output maximised[11]. I will show below that performance pay is such an innovation.

5.2.2 Baseline with Noise - Pre-Reform

Before the academic pay reform, professorial wages only increased with age and did not vary with performance. Academics might still derive utility from being more productive (any form of intrinsic motivation, a greater likelihood to attract research funds from external sources, possible bonuses for C4 professors[12], etc.), but this benefit is not related to an academic's salary. The α in the model captures this non-monetary benefit from own academic output. Furthermore, I assume that an academic's utility when matched does not only depend on his productivity utility $\alpha f(\theta_i \mid \theta_j)$, but also on whether i gets along well with j, the location of the faculty of i and j, etc. This is represented by a noise term ν_{ij} that represents the (dis)utility to i from matching with j. For ease of exposition, let $f_2 = \frac{\partial f(\theta_i|\theta_j)}{\partial \theta_j} = c_2 > 0$ (*assumption* 1). I will make the following assumptions regarding the distribution of the idiosyncratic noise term:

- *assumption* 2.a[13] : ν_{ij} are random and independent draws from a symmetric and

[10] In expectation, if no stable matching exists
[11] If the academic output function exhibits increasing differences
[12] Cf. footnote 1 in the previous chapter.
[13] Choo and Siow (2006) and Siow (2009) assume an extreme value distribution for the noise term in their stochastic Becker model. Furthermore, their noise term is specific for the type of possible match partner only, and does not vary with the specific identity of the potential matching partner. This specific distribution allows them to derive neat quasi-demand and supply equations for agent types using results from McFadden (1974), and, in turn testable implications for the empirical distribution of matches. The identification in this chapter however only derives from the change in pay scheme, which in turn changes the utility from own academic productivity. In order to analyse the effect of performance pay on matching assortativeness, I derive the change in probability that the baseline matching $\tilde{\mu}$ is stable. For this I need a symmetric, mean zero distribution and, later on, will assume a uniform distribution to make the analysis more tractable.

mean zero distribution $g(.)$, with support $\left[-\alpha c_2\left(\bar{\theta}-\underline{\theta}\right), \alpha c_2\left(\bar{\theta}-\underline{\theta}\right)\right]$ and pdf $G(.)$

- *assumption* $2.b$: ν_{ij} is independent from θ_i, θ_j

Proposition 2 (Pre-reform matching): *In the academic job market with noise, a less than maximal positive assortative matching (stable or not) arises with non-zero probability*[14]. *This probability is smaller if complementarities are stronger*[15]. (Proof in Appendix A)

Here, a *less than maximal positive assortative matching* is a matching in which the (average) difference in productivity rank of matched academics active in academia is more than 1. More generally, I will consider a matching μ to be *more positively assortative* than a matching μ' if the average difference in productivity rank of matched academics active in academia is smaller in μ than in μ'.

In the next section I analyse if the introduction of performance pay can reduce the probability with which a less than maximal positive assortative matching arises and thus, if $f\left(\theta_i \mid \theta_j\right)$ exhibits increasing differences, reduce the probability with which total academic output is less than maximal.

5.2.3 Baseline with Noise, Bonuses and Tournaments - Post-Reform

The academic pay reform introduces three measures that change an academic's utility from working in academia. First, the reform introduces performance bonuses that can be paid as wage supplements to attract outside professors or prevent professors from wandering off. Second, after the reform bonuses for research or educational performance can be won in on-the-job tournaments. Third, professors can be paid a supplement from third-party awarded funds for research or teaching projects for the duration of such projects since the reform.[16]

The effect of the on-the-job tournaments on an academic's utility and consequent matching is ambiguous. The tournaments increase the benefit of working with an academic whose productivity is less than one's own, as this would increase the probability that one would

[14]Where, in the case when no stable matching exists, I use the set of absorbing matchings as solution concept (as suggested by Klaus et al. (2010)).

[15]Complementarities are stronger if $\partial f\left(\theta_i \mid \theta_j\right) / \partial \theta_j = c_2$ is larger

[16]As mentioned in Chapter 2, the reform also introduced bonuses for taking management roles or tasks (BMBF 2002). Since these bonuses are not related to academic (research) output, they do not affect the matching process.

win the tournament. But this also means that the tournament increases the disutility from working with someone whose productivity is greater than one's own. Solving for the effect of the tournaments on the stability of the baseline stable matching $\tilde{\mu}$ analytically would require making strong assumptions about the size of the tournament prizes relative to the size of the spillover effects and proportionality constant α. I do not know of stylised facts that could inform any such assumptions and so remain agnostic about the effect possibly going either way.

The attraction bonuses change an academic's utility from working in academia in two ways. Firstly, receiving a greater bonus to become part of a faculty (match with another academic) simply increases the utility an academic derives from being part of that faculty. If the amount of money a faculty (a pair of academics) has to spend on bonuses does not vary in a systematic way with the productivity of the academics, this simply comprises another noise term. A faculty that has a greater budget to spend on attraction bonuses should be able to attract better academics.

Secondly, the prospect of future attraction bonuses should increase the utility people derive from academic productivity, since a greater productivity now increases the chances that an academic will be offered a (higher) attraction bonus to take up a position at a (new) university or a (higher) retention bonus if the academic is already affiliated with a university and that university does not want to lose the academic. Similarly, a greater productivity increases the chances that an academic will continue to receive his attraction or retention bonus if this is not awarded permanently.

The supplements that professors can get paid from third-party awarded funding should also increase the utility academics derive from academic productivity if a greater productivity increases the chances that an academic is awarded such funding. This increase in the utility derived from academic productivity could be captured by an increase in the weighing constant α in the model. In turn, this should increase any academic's preference for matching with a high productivity academic, since working in the same faculty as a high productivity academic increases own academic output if there are positive productivity spillovers.

Proposition 3 (Post-reform matching): *If there are spillovers between academics and if the utility from academic productivity is greater post-reform (i.e. if $\alpha^{post} > \alpha^{pre}$), it is more likely that the maximal positive assortative matching is stable, and less likely that* *any matching that is a less than maximal positive assortative matching is stable. Hence, if there is a stable matching post-reform, it is more likely to be a more positively assortative matching. This effect is stronger when complementarities are larger.* (Proof in Appendix A)

Proposition 3 tells us that performance pay increases the probability that the maximal positive assortative matching is stable if there are positive spillover effects, and hence that academic output is maximised if the academic output function exhibits increasing differences. Moreover, the likelihood that the maximal positive assortative matching is stable again is larger if complementarities are stronger. This implies that we can test whether performance pay increases positive assortative matching when there are positive spillovers by comparing the change in assortativeness upon the introduction of performance pay between fields with different complementarity strength. Fields in which complementarities are stronger should see a larger increase in positive assortativeness if performance pay increases positive assortative matching. I test this hypothesis in the next section.

5.3 Data Description

For this project I use the department level panel data set that I constructed for this project, which encompasses the 89 German public universities between 1999 and 2013. For each department, the panel provides the total number of tenured professors in a given year, the number of new hires into tenured positions, the number of academics already affiliated and in tenured positions, and the number of tenured professors that retire in a given year. The panel further contains the number of new hires who start their first tenured position at a public university (junior hires) and the number of new hires who move from another tenured affiliation (senior hires) and the number of people who leave a tenured position. Apart from information regarding the size of these categories of academics, the department panel also contains productivity variables, most notably average productivity, as well as a proxy for the hiring budget of departments. I give a precise definition and describe the construction of all these variables in the next section.

As explained in the Data chapter, I focus on academics who held a tenured position at a German public university between 1999 and 2013 for the purposes of this research because they are the ones who qualify for performance bonuses in the performance pay scheme and they have a comparable research output.

All in all, the data set used for this chapter covers all departments in all the research active public universities in Germany. The final data set contains 1068 departments in 89 universities over 15 years[17].

5.3.1 Descriptive Statistics

The main dependent variable for the empirical analysis is the average productivity of new hires of a department. I define the average productivity of an academic as the impact factor-weighted number of publications in years t-6 through t-1 divided by the number of years he was academically active in this same period[18]. The average productivity of new hires has a mean[19,20] of 3.68 and a standard deviation of 8.68 (cf. Table 5.1). New hires that start their first tenured position (junior hires) in the sample period have a mean average productivity of 3.67, with a standard deviation of 8.86, while academics who move from one tenured position to another (senior hires) have a mean average productivity of 5.22 with a standard deviation of 14.27. The mean average productivity in year t of academics who were already affiliated with a department in t-1 (labelled "affiliates") is 2.56, while the mean average productivity of those that leave their affiliation at some point in the sample period (the "leavers") is 3.56. The mean of the departmental average of the average productivity of all academics affiliated with a given department (new hires and affiliates) is 2.58.

The most important explanatory variables for this study are the quality of a department and the hiring budget. I use the pre-sample mean of the average productivity of tenured academics affiliated with a department as a measure of the quality of a department[21]. The mean of the department quality variable is 2.22 and the standard deviation is 4.15. I use the number of tenured academics who retire from a certain university in a given year as a proxy for the hiring budget of that university (I explain why I do this and argue that this is a reasonable proxy for the hiring budget in the next section). The average of this

[17] Of the 15 years I have data for, I use 6 for my baseline sample (2001-2006) to abstract from effort effects.

[18] Here, academically active means having passed all necessary qualifications and being active in academia (both research and teaching) and/or publishing articles.

[19] Note that this is a triple mean: it is the average productivity of new hires, averaged over all new hires of a department, averaged over all departments of German universities

[20] This is the mean calculated for new hires not coming from outside Germany and for the sample period, 2001-2006, used for most of the empirical analysis below. The reason for restricting the sample period and set of new hires is discussed in the next section.

[21] This is the average of the average productivity of all academics affiliated with a department in 1999 and the affiliates in 2000

variable is 7.28, with a standard deviation of 6.89.

The average department has 22.77 tenured academics, 21.60 of which are existing affiliates on average, and 1.10 is a new hire. This new hire is more often an academic that starts his first tenured affiliation than one that moves from another tenured affiliation; the average number of junior hires at a department in a given year is 0.78, while the average number of senior hires is only 0.32.

5.4 Empirical Analysis

The main aim of the chapter is to test whether a steeper incentive scheme increases positive assortative matching by productivity. Positive assortative matching can increase through two channels: if higher quality departments are able to hire better academics and if higher quality departments fire lower quality academics. Amongst the hires of a department, we can make a further distinction between junior and senior hires, with junior hires starting their first tenured affiliation and senior hires moving from one tenured affiliation to another. Accordingly, I analyse how the quality of all new hires, junior and senior hires and leavers[22] changes with departmental quality before and after the reform. I first analyse differences in the quality of new hires and leavers before and after the reform for high and low quality departments to establish whether positive assortativeness increases. If performance pay increases positive assortative matching, the difference in the quality of new hires between high and low quality departments should be larger after the reform. Moreover, if this increase is driven by complementarities, the difference in quality of new hires between high and low quality departments should be larger in fields in which complementarities are larger. I therefore analyse differences in changes in positive assortative matching between high and low complementarity fields as a second step. Finally, I control for alternative explanations such as existing pre-trends and systematic differences in hiring budget between departments.

5.4.1 Descriptive Evidence

Before starting the actual analyses, I first need to assess if there is a large influx of academics from outside Germany and, in particular, if the influx changes after the reform. To this

[22] I do not actually observe whether an academic gets fired or leaves out of own free will, so I will refer to academics who (are made to) leave a department as "leavers" in what follows.

end, I show the number of new hires coming from outside Germany as a percentage of total new hires in a given year for the sample period (2001 - 2006) in Figure 5.5.1. The mean percentage of new hires from outside Germany over this period is 6.14. Importantly, there is no marked increase in the percentage of new hires coming from outside Germany after performance pay is implemented in 2005. Nevertheless, I restrict my sample to exclude new hires attracted from outside Germany for the empirical analyses below in order to make sure my results are not driven by a change in the composition of influx or sorting patterns of new hires coming from outside Germany but by a redistribution of academics within Germany[23].

As a first check for changes in assortativeness, I look at the distribution of the average productivity of new hires, junior hires, senior hires and leavers over time at departments of different quality. Panels a, b and c of Figure 5.5.2 show the kernel density plots of the average productivity of new hires in the year before the reform came into effect (2004) and the year after (2006), for departments in, respectively, the top quartile, the second quartile from the top (the third quartile), and the bottom two quartiles of my departmental quality variable - the mean average productivity of affiliates of departments. The mean average productivity of affiliates of departments at the 75th percentile is just larger than the mean average productivity of affiliates averaged over all departments in Germany. I restrict the sample to the year before and after the reform to abstract from changes in effort due to the reform. There is a clear rightward shift of the distribution from pre- to post-reform for the top quartile departments, while there is a slight shift to the left (if any) for departments in the third quartile and no clear shift in the bottom two quartiles. This shows that high rank departments can attract better candidates post-reform.

A similar pattern can be seen for junior and senior hires in Panels a through c of Figures 5.5.3 and 5.5.4 respectively, with the rightward shift for top quartile departments being clearest for junior hires. The leftward shift of the distribution for lower quality departments is most clear in the bottom two quartiles for senior hires. This shows that lower-rank universities are less able to attract higher quality academics post-reform.

Panels a, b and c of Figure 5.5.5 show the pre- and post-reform[24] kernel density plots of

[23]Given that I construct academics' publication records using the set of publications recorded in ISI as having at least one author with a German affiliation, I do not have a meaningful measure of the average productivity of new hires coming from outside Germany in the current data set. I am working on constructing representative publication records for new hires coming from outside Germany in order to analyse their average productivity and sorting patterns in to German departments in the future.

[24]The sample is not restricted to the year before and after the reform only for these plots, since effort

the average productivity of leavers for departments, respectively. There is a clear leftward shift for high rank departments, suggesting lower quality academics (are made to) leave higher rank departments post-reform. For departments in the second highest quartile and bottom half both the mass at 0 as well as at higher average productivities increases post-reform, so very low quality academics (are made to) leave mid-rank departments, while these departments also lose higher quality academics post-reform. Taken together, these patterns are consistent with a redistribution of higher quality academics from lower rank departments to higher rank departments and of lower quality academics from high rank departments to low rank departments or outside academia - in other words, positive assortative matching increases post-reform.

5.4.2 Regression analysis

To formally test whether positive assortative matching increases post-reform, I estimate the following random effects panel data model[25] for department j in field f and year t:

$$\bar{y}^{\{k\}}_{j,f,t} = \beta_0 + \beta_1 \bar{y}^{old}_j + \beta_2 t + \beta_3 post + \beta_4 post \cdot \bar{y}^{old}_j + c_f + u_{jt} \tag{5.4.1}$$

The dependent variable $\bar{y}^{\{k\}}_{j,f,t}$ is the average productivity of $\{k\}$ in department j in field f and year t, where $\{k\}$ denotes new hires, junior hires, senior hires or leavers. The main explanatory variable is $\bar{y}^{old}_j$; the average productivity of the affiliates of department j in the pre-sample years 1999/2000, which I use as a measure of department quality. The variable *post* is zero before the reform (t<2005) and one thereafter and $post \cdot \bar{y}^{old}_j$ is the interaction of this variable with department quality. The c_f are field fixed effects[26].

From the theoretical framework we have that higher quality departments are able to attract better academics on average, so the coefficient on the department quality variable should be positive for regressions with average productivity of new hires, junior hires or senior hires as dependent variable. At the same time, an academic leaving a higher quality department is on average a better academic than one leaving a lower quality department,

changes due to the reform would actually go against the pattern expected for leavers if there is an increase in positive assortative matching.

[25] I estimate a random effects model so that I can estimate the coefficient of departmental quality and use this as a benchmark to compare the interaction of *post* and departmental quality to. The results are robust to estimating the model as a fixed effects model, as can be seen in Table 5.7

[26] The fields distinguished here are the 12 fields along the lines of which I define departments, as explained in the data description section.

so we would expect the coefficient of department quality to be positive in regressions with average productivity of leavers as dependent variable too. Moreover, if the introduction of performance pay increases positive assortative matching, the difference in the productivity of new hires between high quality and low quality departments should be larger post-reform and the interaction of *post* and department quality should be positive too in regressions with average productivity of new hires, junior hires or senior hires as dependent variable. In contrast, we would expect the interaction of *post* and department quality to be negative in a regression with average productivity of leavers as dependent variable if lower quality academics (are made to) leave higher quality departments after the reform.

I restrict the sample used for the regressions to the years 2001 through 2006 for all analyses reported below. This is the period spanning the year before the announcement of the reform to the first year after its implementation. I do so to abstract from changes in effort due to the reform[27] and to minimise any bias introduced by reforms and other events taking place around the same time as the reform[28]. I estimate the specification both with and without a linear time trend t and cluster the standard errors by department.

The results of the regressions are presented in Table 5.2. Throughout the chapter I organise regression results tables in the following way: in columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers is the dependent variable. For each pair of columns, the column labelled 'a' shows the estimation results of the specification without a linear time trend, while a linear time trend is added in column b.

As expected, the coefficient on department quality is positive and significant in all regres-

[27] Restricting the sample to include only the reform year and the year thereafter shuts down the effort channel, since the average productivity variables are defined as the average of the weighted number of publications of an academic over the years t-6 to t-1, so in 2006 I only take publications up until 2005 into account for an academic's productivity measure. It is unlikely that any effort changes in the reform year 2005 immediately affect an academic's publication record in the same year due to e.g. publication lag, but effort changes might start to affect publication records in later years.

[28] The most notable of these is the "Excellence Initiative"; an initiative to boost German research and science through awarding large amounts of funding for projects in either of three categories; research clusters, graduate schools and institutional strategies for top-level research. The first call for proposals for the initiative was given out in 2005, and decided in October 2006, the second round was given out in 2006 and decided in October 2007 (DFG 2014). If, as is likely, higher rank departments are more likely to be awarded funding through the Excellence Initiative, this in itself might make these high rank departments more attractive to high quality academics. This would give rise to a positive coefficient of the rank of a department even if the pay reform does not increase positive assortative matching. I therefore restrict my sample period to end at 2006, so as to prevent this omitted variable bias as much as possible.

sions, so even before the reform academics matched assortatively by productivity. Column 1b shows that if the pre-sample mean of the average number of weighted publications in the previous 6 years of a department's affiliates is higher by 1, the department can hire an academic whose average number of weighted publications in the previous 6 years was 0.33 higher. This number is a bit lower for junior hires, it is higher for senior hires and even higher for leavers. The fact that the coefficient of department quality is highest in the leavers regressions - even higher than in the senior hires regressions - might be because the very best academics leave Germany to start working at top universities abroad, most notably in the U.S.

The coefficients of the interaction of the *post* variable with department quality show that positive assortative matching by productivity increases after the reform. In the regressions with average productivity of new hires and junior hires the interaction of *post* with department quality is positive and significant. The coefficient of this interaction is large: it is 1.2 times the coefficient of department quality in the new hires regression (column 1b) and even 1.8 times the coefficient in the junior hires regression (column 2b). This means that post-reform, the difference in average productivity of new hires between high quality and low quality departments more than doubles, suggesting that departments become more homogenous in terms of average productivity of its affiliates and hence that positive assortative matching increases. This increase in positive assortative matching is driven by junior hires matching more assortatively, a finding that is consistent with the model's predictions if junior hires experience the strongest spillover effects. This would align with the findings in Waldinger (2012) and (2010) that spillover effect are largest for early-career academics [29].

Reassuringly, I observe the opposite pattern in the regressions with the average productivity of leavers as dependent variable: the coefficient of the interaction of *post* with department quality is negative and significant. The coefficient of the interaction term is again sizable; it's absolute size is 54% of the coefficient of department quality. The fact that the interaction is negative means that lower quality academics (are made to) leave higher quality departments post-reform.

Taken together, these results show that positive assortative matching increases signifi-

[29]Waldinger (2012) and (2010) finds that there are positive and significant spillover effects for PhD's, while junior hires in this chapter must have finished a habilitation or equivalent post-doctoral qualification after their PhD and hence are at least six years further in their career.

cantly post-reform and that the increase is large economically. The linear time trend in the b columns is never significant and including the time trend hardly changes the coefficients on the interactions of the *post* variable and department quality, so there is no evidence of a gradual increase in academic quality over time.

5.4.2.1 Baseline Regression

From the theoretical model we have that if performance pay increases an academic's pay-off from academic output, matching is expected to become more positively assortative if there are positive spillovers between academics, and moreover, the increase in positive assortative matching is expected to be larger if complementarities are stronger. Hence one way to test if performance pay increases positive assortative matching is to compare the difference in changes in positive assortativeness in fields with weaker and stronger complementarities, effectively using the strength of complementarity in a field as a measure of treatment strength. If performance pay increases positive assortative matching, the increase should be larger in fields in which complementarities are stronger (i.e. receive a stronger treatment). I test this in the following triple-differences baseline regression :

$$\bar{y}^{\{k\}}_{j,f,t} = \beta_1 \bar{y}^{old}_j + \beta_2 Complementarity_f + \beta_3 Complementarity_f \cdot \bar{y}^{old}_j + \beta_4 t + \beta_5 post$$

$$+\beta_6 post \cdot \bar{y}^{old}_j + \beta_7 post \cdot Complementarity_f + \beta_8 post \cdot Complementarity_f \cdot \bar{y}^{old}_j + c_f + u_{jt} \tag{5.4.2}$$

Here, $Complementarity_f$ is the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000), which I use as a proxy for the strength of complementarities in a field. All other variables are as defined above. I use the average number of authors on a paper in a field as a proxy for the strength of complementarities, since the larger this number, the more prevalent is collaboration and the greater the opportunity for spillovers[30]. By this measure, complementarities vary widely across fields. The average number of coauthors is 3.02 per paper, ranging from 1.02 in the field of theology

[30]Indeed, Azoulay et al. (2010) and Waldinger (2012) find positive spillover effects amongst co-authors while Agrawal et al. (2014) find positive spillover effects for co-located colleagues who work on related subjects.

to 4.63 in the field of medicine and 6.43 in the field of physics, mathematics and computer science.

The estimation results of regression 5.4.2 for new hires, junior hires, senior hires and leavers form the core results of this chapter and can be found in Table 5.3. Columns 1a through 3b show that the triple interaction of the *post* variable with department quality and field complementarity strength is positive and significant for all new hires, junior hires and senior hires. This means that the difference in average productivity of new hires between high quality and low quality departments is larger post-reform in fields in which complementarities are stronger, so the increase in positive assortativeness after the reform is larger in high complementarity fields. As explained above this is consistent with performance pay increasing positive assortativeness. Academics in high complementarity fields face stronger incentives to match positive assortatively and if pay for performance increases the pay-off from academic output, the incentive to match positive assortatively increases more in high complementarity fields, so we should see a larger increase in positive assortative matching in those fields when performance pay is introduced. The positive and significant triple interaction of the *post* variable with department quality and field complementarity strength is evidence of this.

The interaction of departmental quality and field complementarity strength is positive and significant for all new hires and junior hires too (cf. columns 1a through 2b). This means that the difference between the average productivity of new hires between high quality and low quality departments is larger in high complementarity fields even before the reform, so matching is more positive assortative in fields with a larger complementarity measure. This is reassuring, since if there are larger spillovers in fields for which my measure of field complementarity strength is larger we should see that matching is more positive assortative in those fields. The effect is driven by junior hires matching more positive assortatively in high complementarity fields, a finding that is consistent with junior hires experiencing stronger spillover effects and that aligns with findings in Waldinger (2012) and (2010) that spillover effects are largest for early-career academics.

Finally, the positive and significant interaction of post and field complementarity strength in Columns 4a and 4b shows that the difference in average productivity of leavers between fields with high and low complementarity is larger after the reform, so more productive academics change affiliation in high complementarity fields. This too is consistent with performance pay increasing positive assortativeness, since academics in high complementarity fields stand to gain most from matching more positive assortatively and so higher productivity academics in these fields are more likely to move to a high(er) quality department when performance pay increases pay-off from academic output.

5.4.2.2 Alternative Explanations

In the previous sections I have shown that matching becomes more positive assortative after the reform and, moreover, that the increase in positive assortative matching is larger in fields in which complementarities are stronger. This is consistent with performance pay increasing positive assortativeness. In this section I want to rule out alternative explanations that might have caused assortative matching to increase and, specifically, more so in high complementarity fields.

5.4.2.3 Pre-existing Trends

First, I test for pre-existing trends. I do so by adding a placebo-post dummy for the year before the reform was implemented (2004) and interactions with this dummy to the baseline specification in the following way:

$$\bar{y}_{j,f,t}^{\{k\}} = \beta_1 \bar{y}_j^{old} + \beta_2 Complementarity_f + \beta_3 Complementarity_f . \bar{y}_j^{old} + \beta_4 t + \beta_5 post'04 + \beta_6 post$$

$$+\beta_7 post'04 . \bar{y}_j^{old} + \beta_8 post . \bar{y}_j^{old} + \beta_9 post'04 . Complementarity_f + \beta_{10} post . Complementarity_f$$

$$+\beta_{11} post'04 . Complementarity_f . \bar{y}_j^{old} + \beta_{12} post . Complementarity_f . \bar{y}_j^{old} + c_f + u_{jt} \quad (5.4.3)$$

The variable $post'04$ indicates a placebo dummy that is 0 for the years before 2004 and 1 otherwise. All other variables and specifications are as in the baseline regression (equation 5.4.2). If there are pre-existing trends towards a greater increase in positive assortative matching in high complementarity fields even before the introduction of performance pay, due to e.g. anticipation of the reform[31], the placebo-post dummy triple interaction with department quality and field complementarity strength should be positive and significant and the triple interaction of the *post* variable[32] with department quality and field complementarity strength smaller.

Table 5.4 shows the estimation results of the specification in 5.4.3. In the all new hires and junior hires regressions the triple interaction of the $post'05$ variable with the department quality and field complementarity strength remains positive and significant and the coefficient is of similar size as in the baseline regression (Table 5.3), especially for junior hires. Moreover, the coefficients of the placebo-post dummy triple interaction are not significant in any of these regressions, and smaller than those of the triple interaction of $post'05$ with department quality and field complementarity strength. Hence there is no evidence of a pre-existing trend for all new hires and junior hires.

The results for senior hires are not robust to controlling for pre-trends: the triple interaction of $post'05$ with department quality and field complementarity strength in the senior hires regression is no longer significant and the coefficient is smaller than in the baseline regression and of similar size as that of the placebo-post dummy triple interaction. In the leavers regression, the interaction of $post'05$ with field complementarity strength loses significance, though the size of the coefficient increases compared to the baseline and the coefficient of the placebo-post dummy triple interaction has the opposite sign.

I also estimate an extended specification that includes a full set of placebo-post dummies

:

$$\bar{y}^{\{k\}}_{j,f,t} = \beta_1 \bar{y}^{old}_j + \beta_2 Complementarity_f + \beta_3 Complementarity_f . \bar{y}^{old}_j + \beta_4 t + \beta_5 post'0\{l\} + \beta_6 post$$

$$+ \beta_7 post'0\{l\} . \bar{y}^{old}_j + \beta_8 post . \bar{y}^{old}_j + \beta_9 post'0\{l\} . Complementarity_f + \beta_{10} post . Complementarity_f$$

[31]Note that other events that would cause such a pre-existing trend would have to be very particular, e.g. an (relatively larger) increase in research funding for high quality departments in high complementarity fields.

[32]Labelled $post'05$ in Table 5.4 for extra clarity.

$$+\beta_{11} post'0\{l\}.Complementarity_f.\bar{y}_j^{old} + \beta_{12} post.Complementarity_f.\bar{y}_j^{old} + c_f + u_{jt}, l \in [2,3,4] \tag{5.4.4}$$

Here $post'0\{l\}$ indicates placebo dummies that are 0 for the years before 2002, 2003 and 2004, respectively, and 1 otherwise. All other variables and specifications are as before. The estimation results of this specification are presented in Table 5.5 and the coefficient estimates and 90% confidence intervals of the triple interactions are depicted graphically in Figure 5.5.6 for the new hires regression (Panel a) and junior hires regression (Panel b) respectively. The results of these regressions are very similar to the pre-trend regressions reported in Table 5.4; the triple interactions of $post'05$ with department quality and field complementarity strength remain positive and significant for all new hires and junior hires and the coefficients have similar sizes as in the baseline regression (Table 5.3), especially junior hires. Moreover, none of the placebo-post dummy triple interactions are consistently positive and significant in these regressions, and the coefficients are generally smaller or even have the opposite sign than those of the triple interaction of *post* with department quality and field complementarity strength. Hence I do not find evidence of a pre-existing trend that can explain the larger increase in positive assortative matching in fields with stronger complementarities. It is the robust larger increase in positive assortative matching by junior hires in fields with stronger complementarities that drives the differential increase in positive assortativeness.

5.4.2.4 Hiring Budget

The above results do not provide sufficient evidence to prove that the introduction of performance pay causes the increase in positive assortative matching. An alternative explanation for the increase in positive assortative matching could be that higher rank departments in fields with stronger complementarities have a greater hiring budget and therefore can 'afford' better new hires. Before the reform, German universities did not have autonomy regarding the salary offered to a candidate since any professor was paid according to his age[33] Post-reform however, universities can offer bonuses to attract academics or prevent

[33] Before the reform C4-Professors could be awarded pay supplements in special cases. However, this concerned only few professors and was the responsibility of the respective state's ministry (Handel 2005). See also footnote 1 in the previous chapter.

affiliates from moving to another university[34]. A difference in hiring budget can therefore make a real difference in the quality of new hires a university can attract post-reform. In order to distinguish between better departments in high complementarity fields simply 'buying' better new hires post-reform and positive assortative matching increasing more in high complementarity fields due to the introduction of performance pay, I control for the hiring budget of a department in the following series of regressions.

I do not observe a department's hiring budget. Moreover, the actual, instantaneous hiring budget of a department would likely not be a very good control for the budget in a regression of new hire quality because of omitted variable bias concerns; for instance if management actively attempts to increase quality by both forcing more lower quality academics out and using the budget thus freed up to hire better academics. I therefore use the number of academics who retire from a university as a proxy for its hiring budget. This measure is historically determined, allaying endogeneity concerns, and is a source of variation of the hiring budget as I will argue next.

As discussed, the academic pay reform in Germany includes a requirement that the average professorial pay at the federal ("Bund") and state ("Land") level remain at the respective levels before the reform (BMBF 2002). Both the personnel budget and the number of professors a university can employ is determined by the ministry of education of the respective state (in the "Stellenplan") and this does not vary much from year to year (Kaiser et al. 2002, Jongbloed 2009). Combined with the fact that under the old, age-related pay system academics close to retirement earn the highest salary, this means that a university from which many academics retire in a given year has a larger hiring budget.

German law stipulates that academics retire at the age of 65 (Mohr 2007, *Bundesbeamtengesetz* 1985). Because a professor who is about to retire will turn 65 in the course of a year and because positions are likely not immediately vacated, let alone filled, I will use the number of tenured professors that turn 66 in a given year as a proxy for the hiring budget of a given university (Pritchard 2006). Given that the age of retirement is mandated by law, the number of academics who retire from a department in a given year is historically determined and should not be correlated with contemporaneous factors[35].

[34] See Handel (2005) for a comprehensive overview of the degrees of autonomy of higher education institutes due to the reform in the different German states

[35] The retirement year can be extended beyond the 65th year upon the request of the academic and if this is in the interest of the university. Such an extension can be granted for a limited time only of up to a year

The hiring budget variable I use here is the total number of academics retiring from a university, not the specific department. Since the late nineties, more and more German states have introduced more lump sum budgets ("Globalhaushalte") that give universities a greater flexibility to allocate funds and positions across departments and institutions as well as over time (Vossensteyn and Jongbloed 1998, Kaiser et al. 2002, *"Personalmittelbudgetierung, Empfehlungen zu ihrer Ausgestaltung"* 2008). Moreover, in a study of personnel budgeting at German universities performed by a task force of the German university chancellors, it was found that in all the universities studied, the allocation of performance pay bonuses is decided at the university level (*"Personalmittelbudgetierung, Empfehlungen zu ihrer Ausgestaltung"* 2008). I will therefore use the hiring budget - and in particular the portion available for performance pay - defined at the university level for the regressions below. I will however also present estimation results of regressions with hiring budget defined at department level[36] as a robustness check later on.

The regressions in which I control for hiring budget have the following specification:

$$\bar{y}^{\{k\}}_{j,f,t} = \beta_1 \bar{y}^{old}_j + \beta_2 Complementarity_f + \beta_3 Complementarity_f \centerdot \bar{y}^{old}_j + \beta_4 B_{j,t} + \beta_5 t + \beta_6 post$$

$$+\beta_7 post \centerdot \bar{y}^{old}_j + \beta_8 post \centerdot Complementarity_f + \beta_9 post \centerdot Complementarity_f \centerdot \bar{y}^{old}_j + \beta_{10} post \centerdot B_{j,t} + c_f + u_{jt} \tag{5.4.5}$$

The variable $B_{j,(t,t-1)}$ is the number of professors that retire (turn 66) between $t-1$ and t from the university to which department j belongs. As explained above this is my proxy for the hiring budget of department j in year t. The coefficient on its interaction with the *post* variable should be positive in regressions with average productivity of new hires, junior hires or senior hires if departments that have a larger hiring budget are able to attract better academics post-reform. All other variables and specifications are as before[37].

every time, with the total not exceeding the 68th year of the academic (*Bundesbeamtengesetz* 1985). Such extensions do not seem to occur frequently, and indeed German academics who have reached the age of 65 but wish to continue working in academia have been known to emigrate, most notably to the US (Mohr 2007).

[36] In particular, the number of affiliates of a specific department that turn 66 in a given year

[37] I also estimated a specification in which I additionally control for the total number of academics retiring from all German universities in year t. I use this variable as a proxy for the number of vacancies in the German university system in a given year and include it to control for tightness of the academic job market. In a year in which many academics retire, there are likely more job openings and hence a

The estimation results of this specification are reported in Table 5.6. Firstly, note that the triple interactions of *post* with department quality and field complementarity strength in the regressions for all new hires, senior hires and junior hires remain positive and significant after controlling for hiring budget (cf. columns 1a through 3b). Moreover, comparing the results from the baseline regressions in Table 5.3 with the results in Table 5.6 shows that adding hiring budget as a control barely changes the coefficients of the triple interactions of *post* with department quality and field complementarity strength. If higher rank departments in high complementarity fields would have a larger hiring budget post-reform, and if this would drive the differential increase in positive assortative matching in fields with different complementarity strength evidenced by the baseline regressions, the coefficient on the triple interactions of *post* with department quality and field complementarity strength should decrease when controlling for hiring budget (the formerly omitted variable), but I do not find evidence of this.

Secondly, columns 1a and b show that the interaction of the *post* variable with hiring budget is positive and significant in the regressions with the average productivity of new hires as dependent variable. The size of the coefficient means that if, ceteris paribus, one more academic retires from a university after the reform, its departments can attract an academic with 0.10 more weighted publications on average. This coefficient is small compared to that of the triple interaction of the *post* variable with department quality and field complementarity strength; it is only 25% of the size of the latter. The positive effect of hiring budget on new hire quality after the reform is driven by departments with larger hiring budgets being able to attract higher average productivity senior hires: the interaction of the *post* variable with hiring budget is positive and significant in the senior hires regressions (columns 3a and b) but insignificant in the junior hires regressions (columns 2a and b). This discrepancy between the effect of hiring budget on senior and junior hire quality can be explained by the difference in (base) wages between the age-related and the performance pay system.

The base wage in the performance pay system is lower than the wage at all but the lowest ages in the equivalent age-related pay level (Hochschullehrerbund 2009). Moreover, any offer accepted after the reform (or renegotiation of a current position) results in an academic to be paid under the new, performance based pay scheme (Detmer and Preissler

greater demand for new hires. This might force departments to hire academics of lower quality than they would in a year in which demand is lower. Results are robust to inclusion of this variable.

2005). It seems reasonable to assume that an academic would not want to take a pay cut when taking on a new position (cf. Detmer and Preissler 2006)[38]. An academic who already has a tenured affiliation before the reform will then only consider an offer that at least matches his current age-related pay through the payment of a large enough attraction bonus. Thus, academics who already have a tenured affiliation under the old pay system need to be compensated for the lower base wage in the new system. There is no such need to compensate junior hires since they don't experience a drop in base wage (Detmer and Preissler 2004, 2006). The finding that the interaction of the *post* variable with hiring budget is not significant in the junior hires regressions, but positive and significant in the senior hires regression is consistent with this difference in the need to compensate for a drop in base wage between junior and senior hires.

Thirdly, columns 4a and b show that the interaction of the *post* variable with hiring budget is also positive and significant in the leavers regressions. This suggests that a greater hiring budget allows departments to prevent medium productivity academics from leaving by offering a retention bonus, but not top-level academics. In turn, this is consistent with positive spillover effects. If academic output increases in partner's productivity, a high productivity academic is more likely to receive offers from higher quality departments than a medium productivity academic. A high productivity academic then needs to be paid a larger retention bonus to prevent him from leaving than a lower productivity academic at a department of the same quality, since the high productivity academic needs to be compensated for a larger difference in academic output and corresponding expected future performance pay than a lower productivity academic. If spillovers are sufficiently large, departments might simply not have a large enough hiring budget on average to compensate high average academics, but their budget might be large enough to compensate and retain mid-level academics. This would show up as a positive and significant interaction of the *post* variable with hiring budget, which is what I find here[39].

[38]This assumption holds if academics are risk averse or discount future pay. If an academic is risk neutral, as is currently assumed in the model, he should prefer a higher rank department that offers a lower pay than his current age-related wage, if working at the higher rank department increases his productivity enough (through productivity spillovers) that (the present value of) the sum of expected future attraction bonuses and supplements from third-party awarded funding and the base wage is at least as large as (the present value of) his age-related wage. If he is risk-averse, (the present value of) the sum of expected future attraction bonuses and supplements and the base wage have to be larger than (the present value of) his age-related wage for an academic to change affiliation. Similarly, if future pay (utility) is discounted, the present value of expected future attraction bonuses and supplements plus base wage have to be larger than the present discounted value of his age-related wage for an academic to change affiliation. In the limit, academics would not accept a cut in current wage.

[39]Note that supermodularity of the academic output function could also explain why the interaction of the

5.4.2.5 Robustness Checks

Finally, I do a number of robustness checks, the results of which are reported in Table 5.7. I only report regressions with the average productivity of all new hires as dependent variable. The results are similar with average productivity of junior hires, senior hires and leavers as dependent variable. In the first column I estimate the baseline specification (equation 5.4.2) as a fixed effects panel data model[40]. The estimation results barely change compared to the baseline regression in Table 5.3, columns 1a and 1b. Importantly, the triple interaction of the *post* variable with department quality and field complementarity strength remains positive and significant and the coefficient has a similar size as in the baseline regressions in Table 5.3. This is reassuring, since if competition for personnel funds within a university or between universities in a state would drive the result, we should see a smaller coefficient of the triple interaction in the fixed effects model.

Column 2 reports the estimation results of the random effects baseline model with year fixed effects. This specification controls more flexibly for any changes in average productivity of new hires over time than the baseline specification with a linear time trend. The estimation results are however very similar to the baseline results in Table 5.3, and importantly, the triple interaction is virtually identical.

In column 3a and 3b I show the results for the baseline regression estimated using an extended, balanced panel, spanning the years 2001 to 2009 (from 4 years before implementation of the reform to 4 years after). The triple interaction of the *post* variable with department quality and field complementarity is positive and significant here too, though the size is smaller than in the baseline regression in Table 5.3. This could be explained by the large funding waves for research and academic education that started at the end of 2006 and 2007 (the "Excellence Initiative"). In particular, the initiative awarded large sums of money for projects in either of three categories; research clusters, graduate schools and institutional strategies for top-level research (DFG 2014). Through this initiative,

post variable with hiring budget is positive and significant for senior hires but not for junior hires. The (differential) increase in positive assortative matching is most robust for junior hires, suggesting that they face the strongest spillovers. If these spillovers do not just take the form of the academic output function increasing in a colleague's productivity, but also of supermodularity, high productivity junior candidates would need to be compensated (much) more for not joining a high quality department that will have them than lower productivity junior candidates. If the cross-derivative is large enough, a larger hiring budget might simply not be able to compensate high productivity junior candidates for not joining the highest quality department they can go to, and hence hiring budget would not have explanatory power in the junior hires regression.

[40] Given that the unit of observation is department j, the FE model controls for individual department fixed effects.

(high quality) departments in all fields - not just those with strong complementarities - were awarded sizable funds that they could use to i.a. attract high productivity academics. This should cause the triple interaction to shrink, which is what we see in columns 3a and 3b.

Lastly, I re-estimate the hiring budget-augmented baseline specification (equation 5.4.5) using a hiring budget variable defined at the departmental level. As mentioned in the previous section, the hiring budget variable I use above is the total number of academics retiring from a university, not the specific department. A differential inequality in the distribution of this budget over a university's departments of different quality and across fields with different complementarity strength would bias the coefficient of the triple interaction of the *post* variable with department quality and field complementarity strength. In particular, if a greater portion of the budget is available to higher rank departments in high complementarity fields, this could cause the triple interaction in the baseline regression to be positive. The regressions in table 5.6 did not show evidence that this was a concern, because controlling for hiring budget (at the university level) did not affect the size of the coefficient of the triple interaction. Another way to test whether better departments in high complementarity fields having a larger hiring budget causes the differential increase in positive assortative matching I find in the baseline regressions, is to run the baseline regression controlling for hiring budget defined at the department level. If I use the number of academics who retire from a given department in a given year as hiring budget variable, any deviation from the distribution of the university hiring budget that is proportional to the number of retirees from departments will be uncorrelated with this alternative hiring budget variable. If better departments in high complementarity fields systematically receive a larger share of the university hiring budget, the departmental hiring budget variable should not have explanatory power. Columns 4a and 4b show that the departmental hiring budget variable does have explanatory power post-reform for the quality of all new hires; the interaction of *post* with departmental budget is positive and significant. This shows that departments with a larger departmental hiring budget are able to attract better new hires after the reform, which is consistent with a distribution of a university's hiring budget across its departments that is proportional to a department's number of retirees in a given year. Furthermore, comparing the coefficient of the triple interaction of the *post* variable with department quality and field complementarity strength in Tables 5.3 and 5.7 shows that they are very similar to those in the baseline regressions. Hence I do not find evidence that it is better departments that have larger hiring budgets that drives the (differential) increase in positive assortative matching.

5.5 Conclusion

This chapter studies the effect of performance related pay on matching assortativeness, and provides empirical evidence consistent with performance pay increasing positive assortative matching. In order to study the effect of performance pay on matching assortativeness, I use the introduction of performance pay in German academia as a natural experiment and a data set comprising affiliation and productivity data of the universe of German academics, which I constructed for this research project. The combination of the nationwide introduction of performance pay in an entire sector, and a data set that encompasses affiliation and productivity information on everyone working in that sector allows for studying the effect of the reform on matching assortativeness, as opposed to for instance sorting into a particular pay scheme (if the reform had not been nation- and sector-wide).

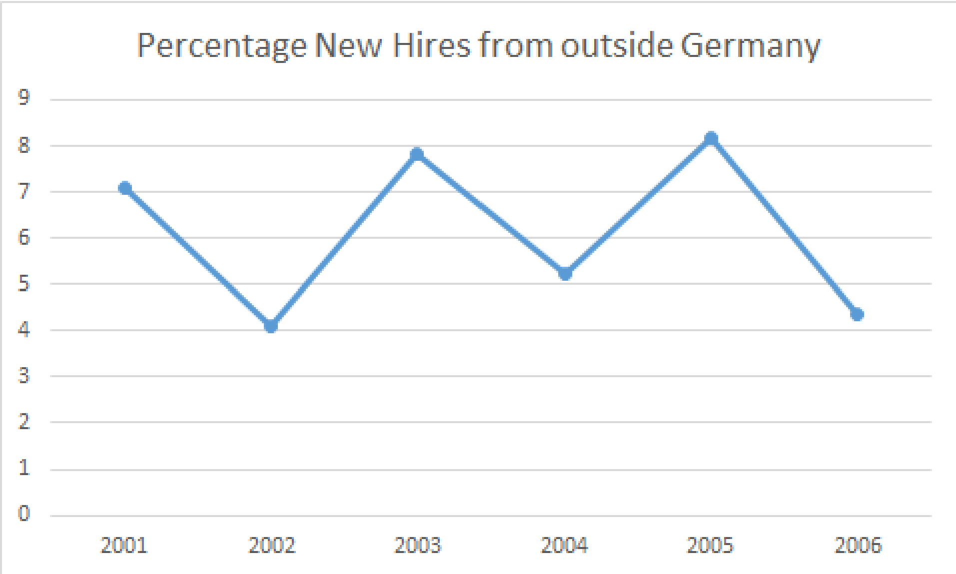
The chapter presents a simple stochastic one-to-one coalition formation model that makes precise the effect of performance pay on matching assortativeness. In the model, an academic's utility comprises a systematic component that depends positively on own and matching partner's productivity, representing positive productivity spillovers, and an idiosyncratic component that solely depends on matching partner identity. The idiosyncratic component represents personal preferences regarding colleagues and place of work and constitutes a friction that causes matching to become less positively assortative. If the academic production function exhibits increasing differences, this implies a decrease in total output. The introduction of performance pay then causes matching to become more positive assortative by productivity. Moreover, this effect is stronger if complementarities are stronger. This result also implies that, if the academic production function exhibits increasing differences in departmental colleagues' productivity so that the maximum total academic output is unique, performance pay increases the probability that total academic output is maximised.

I test the hypothesis that performance pay increases positive assortative matching and that this increase is larger if complementarities are stronger by studying the different channels through which departmental composition may change: hiring - both junior and senior - and firing. First, I find that the difference in average productivity of new hires between high and low quality departments is larger, while the difference in average productivity of leavers between high and low quality departments is smaller after the reform. Hence positive assortative matching increases after the reform. This increase is economically large. Secondly, I estimate whether this increase is larger in high complementarity fields in a difference-in-differences framework in which the strength of complementarities is effectively used as treatment intensity. I find that the difference in average productivity of new hires between high quality and low quality departments is larger post-reform in fields in which complementarities are stronger. The increase in positive assortativeness after the reform is thus larger in high complementarity fields. This is consistent with performance pay increasing positive assortativeness. This result is robust to controlling for alternative explanations such as pre-existing trends and hiring budget, where I use the number of retirees as a plausibly exogenous proxy for this budget.

The study of the effect of performance related pay on matching assortativeness is relevant for two reasons. If there are positive productivity spillovers and if these are such that the productivity of highly productive employees increases when they work with highly productive people and, moreover, if this increase in productivity is larger for highly productive people, then clustering high productivity individuals together increases total output. If welfare depends on total output, this would be welfare-improving. Secondly, the distribution of productive agents (and output) might also directly affect welfare if we care about providing good education to all people across the country for instance. In this case, a more concentrated and hence less equal distribution of high productivity agents actually decreases welfare. This chapter aims to shed some light on the effect of performance pay on matching assortativeness, and, by showing that matching assortativeness increases, finds that incentive scheme reforms may affect welfare in more ways than by increasing individual output only. Academia is a relevant and interesting setting for this study, since the organisation of human capital is of primary importance for knowledge creation and knowledge creation, in turn, is particularly important for innovation and growth.

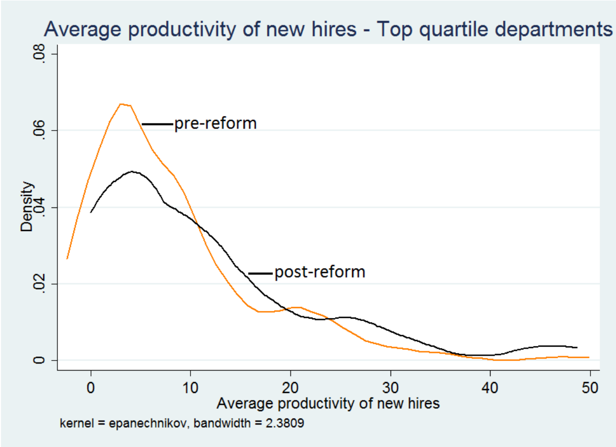
There is a great number of research trajectories that can be taken next. Firstly, it would be very interesting to quantify spillover effects in academia and, in particular, determine whether the academic output function is supermodular. Secondly, estimating the distributional impact of changes in assortativeness on i.a. educational outcomes and university-business collaborations (patents) would be important to obtain a more complete picture of the possible welfare implications of performance pay and changes in assortativeness. Finally, it would be worthwhile to focus on other dimensions of academic output and investigate whether matching becomes more positively assortative by these dimensions as well and whether any such increase is in fact due to the introduction of performance related pay. Results might well be different for dimensions such as student outcomes, as they are the product of the effort of multiple academics, making performance pay based on student outcomes a team-based incentive. Studying the differences in the effects of performance pay along these lines might add to our understanding of the underlying mechanisms that cause performance pay to affect assortativeness and, through it, welfare.

Figure 5.5.1: Percentage New Hires from Outside Germany



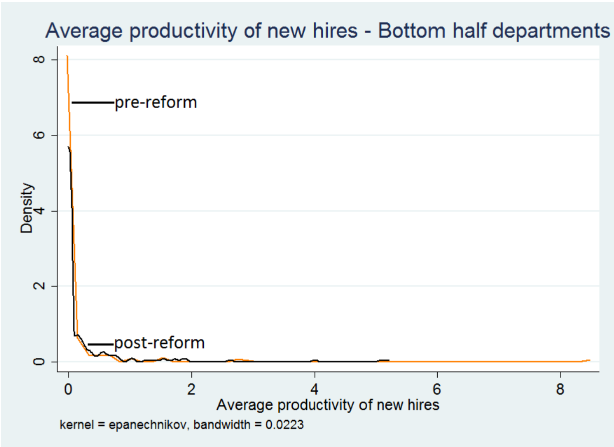
The figure above shows the number of new hires attracted from outside Germany as a percentage of the total number of new hires in a given year for the years 2001-2006.

Figure 5.5.2: Average Productivity of New Hires



(a)

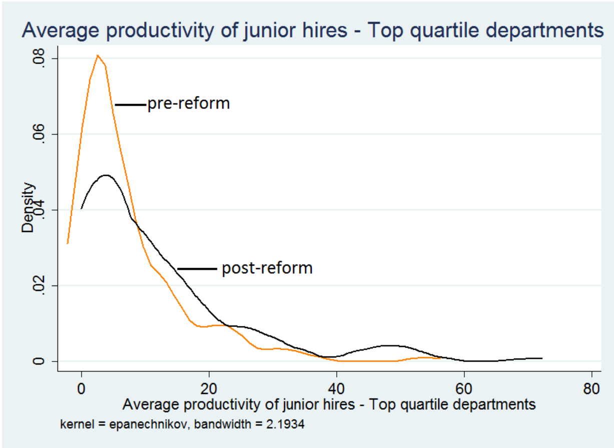
(b)



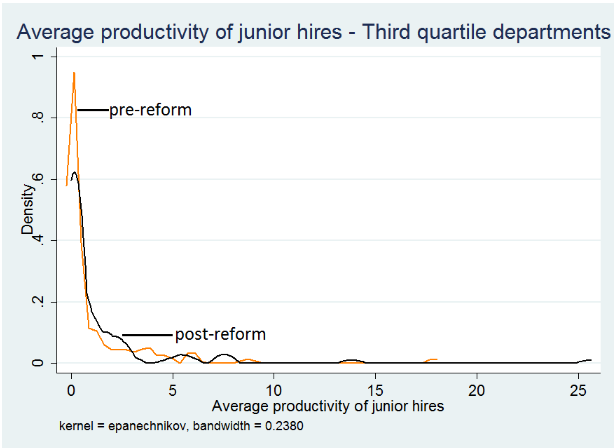
(c)

The above figures are kernel density plots of the average productivity of new hires in the year before the reform came into effect (2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department quality variable. The department quality variable used is the mean average productivity of affiliates of the department. The average productivity of an academic is the average impact-factor weighted number of publications in t-6 to t-1.

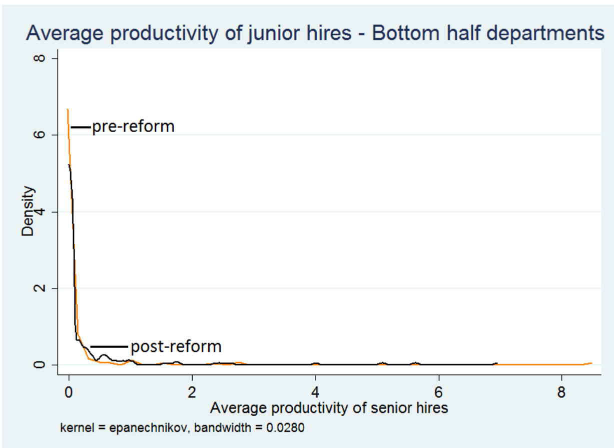
Figure 5.5.3: Average Productivity of Junior Hires



(a)



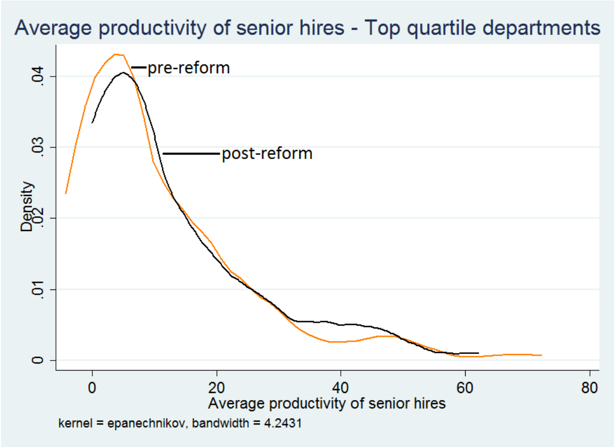
(b)



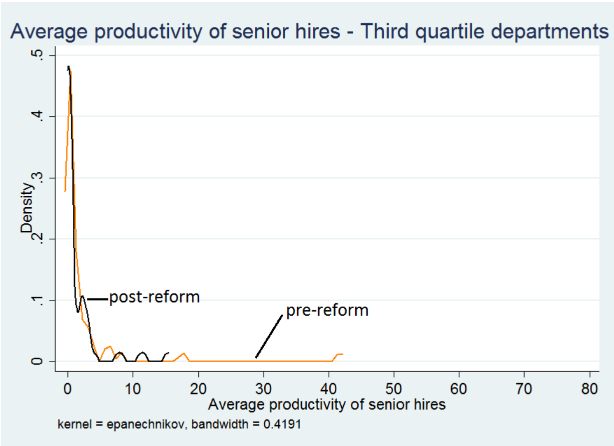
(c)

The above figures are kernel density plots of the average productivity of junior hires in the year before the reform came into effect 2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department quality variable. The department quality variable used is the mean average productivity of affiliates of the department. The average productivity of an academic is the average impact-factor weighted number of publications in t-6 to t-1.

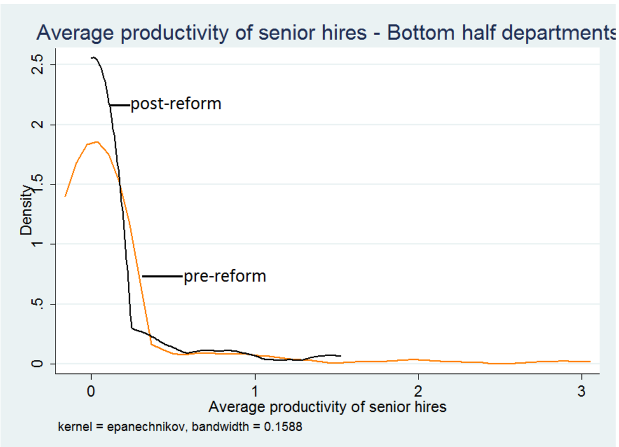
Figure 5.5.4: Average Productivity of Senior Hires



(a)



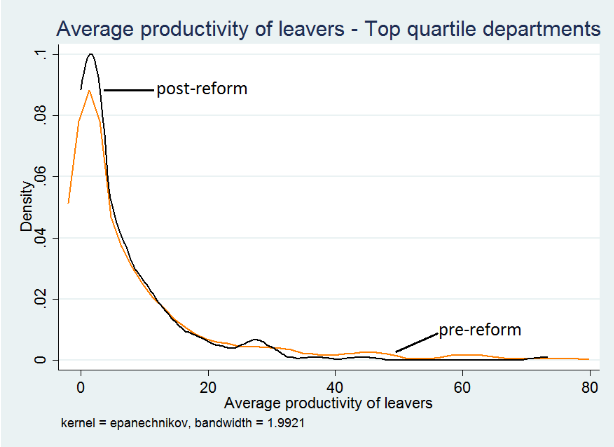
(b)



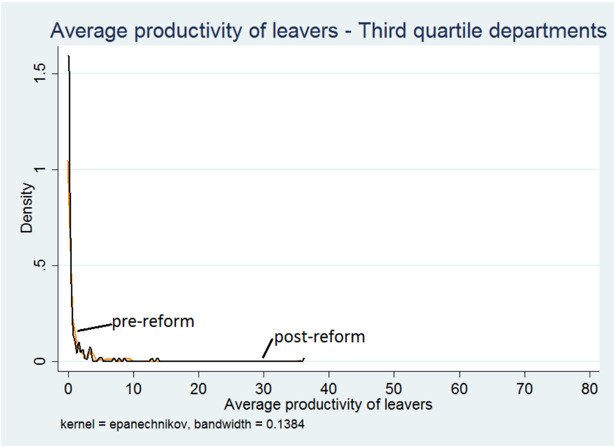
(c)

The above figures are kernel density plots of the average productivity of junior hires in the year before the reform came into effect 2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department quality variable. The department quality variable used is the mean average productivity of affiliates of the department. The average productivity of an academic is the average impact-factor weighted number of publications in t-6 to t-1.

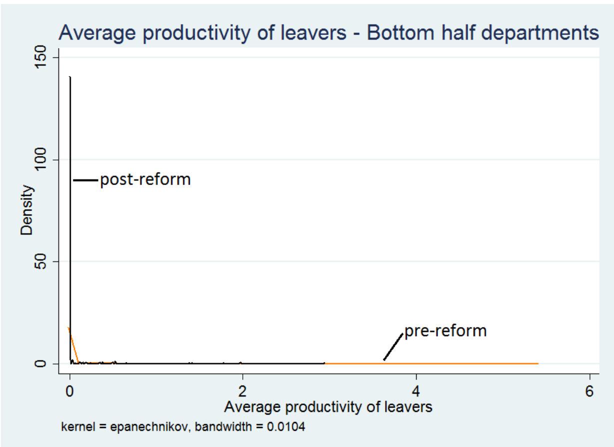
Figure 5.5.5: Average Productivity of Leavers



(a)



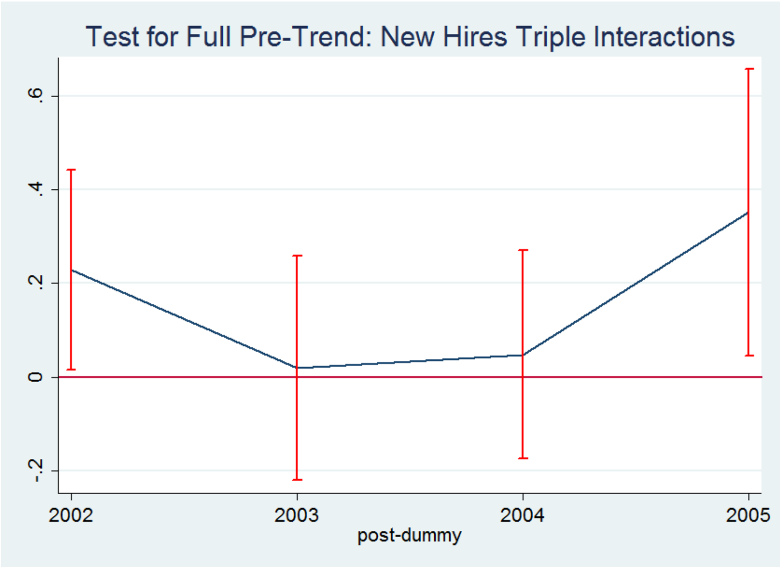
(b)



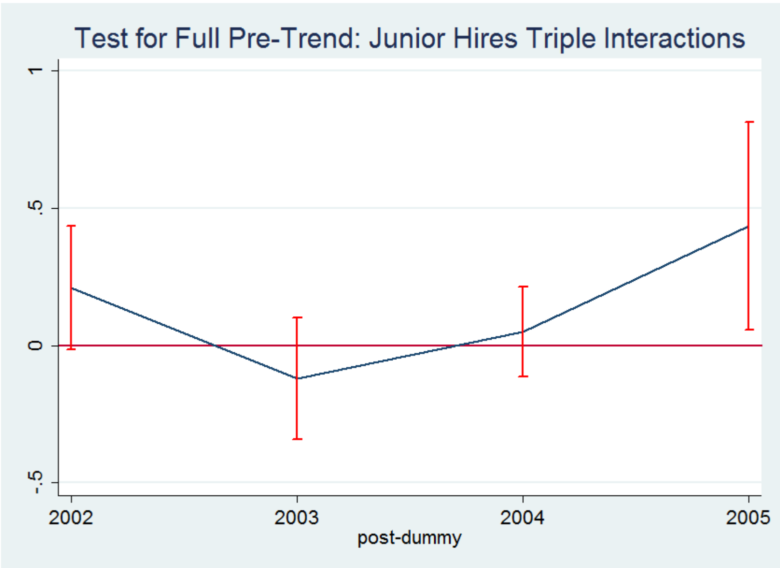
(c)

The above figures are kernel density plots of the average productivity of academics who leave a department in the years before the reform came into effect (2001-2004) and the years after (2005-2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department quality variable. The department quality variable used is the mean average productivity of affiliates of the department. The average productivity of an academic is the average impact-factor weighted number of publications in t-6 to t-1.

Figure 5.5.6: Tests for Full Pre-Trend



(a)



(b)

The above figures depict the coefficient estimates and 90% confidence intervals of the triple interactions of a full set of post-dummies (post'02, post'03, post'04 and post'05) with departmental quality and field complementarity strength in regressions of average productivity of new hires and junior hires respectively on departmental quality and field complementarity strength, a full set of post-dummies, their interactions with departmental quality and field complementarity strength as well as a full set of triple interactions, a linear time trend and field fixed effects. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. The average productivity of an academic is calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Post'0# is 0 before 200# and 1 thereafter. Standard errors are robust, clustered by department.

Table 5.1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Average productivity	5270	2.581	5.045	0	53.957
Average productivity of affiliates	5238	2.556	4.982	0	53.957
Average productivity of new hires	2714	3.682	8.678	0	129.315
Average productivity of junior hires	2313	3.668	8.857	0	144.170
Average productivity of senior hires	1357	5.219	14.272	0	243.572
Average productivity of leavers	2011	3.556	12.234	0	211.211
Department Quality	5166	2.223	4.149	0	35.496
Size	6408	22.766	32.076	0	319
Number of affiliates	6408	21.596	30.669	0	315
Number of new hires	6408	1.101	1.965	0	29
Number of junior hires	6408	0.783	1.586	0	29
Number of senior hires	6408	0.318	0.733	0	7
Number of leavers	6408	0.636	1.449	0	27
Budget (University Level)	6408	7.279	6.889	0	39
Budget (Department Level)	6408	0.679	1.254	0	12
Total number of retirees	6408	797.167	67.113	711	868

Table 5.2: Check for Increase in Positive Assortativeness

	All new hires		Junior hires		Senior hires		Leavers	
	1_a	**1_b**	**2_a**	**2_b**	**3_a**	**3_b**	**4_a**	**4_b**
Department Quality	**0.334*****	**0.334*****	**0.275*****	**0.275*****	**0.441***	**0.447***	**1.130*****	**1.131*****
	(0.091)	**(0.091)**	**(0.087)**	**(0.087)**	**(0.233)**	**(0.234)**	**(0.248)**	**(0.249)**
Linear Time Trend		0.056		-0.092		0.612		0.176
		(0.123)		(0.128)		(0.405)		(0.271)
Post '05	-0.162	-0.330	0.148	0.425	**-1.092***	**-2.895****	**0.950****	0.471
	(0.245)	(0.392)	(0.284)	(0.433)	**(0.599)**	**(1.201)**	**(0.386)**	(0.723)
Post '05 * Department Quality	**0.410*****	**0.410*****	**0.507*****	**0.507*****	0.333	0.332	**-0.606*****	**-0.607*****
	(0.124)	**(0.124)**	**(0.140)**	**(0.140)**	(0.348)	(0.349)	**(0.196)**	**(0.197)**
N	2673	2673	2280	2280	1347	1347	1758	1758

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. In columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Post is 0 before 2005 and 1 thereafter. All regressions contain field fixed effects. Standard errors are robust, clustered by department.

Table 5.3: Baseline Regression

	All new hires		Junior hires		Senior hires		Leavers	
	1_a	**1_b**	**2_a**	**2_b**	**3_a**	**3_b**	**4_a**	**4_b**
Department Quality	-0.227	-0.228	**-0.579****	**-0.578****	**1.156****	**1.155****	-0.247	-0.244
	(0.239)	(0.240)	**(0.254)**	**(0.253)**	**(0.566)**	**(0.564)**	(0.535)	(0.535)
Complementarity	**1.371*****	**1.367*****	**1.271****	**1.279****	**1.567*****	**1.616*****	0.555	0.553
	(0.437)	**(0.440)**	**(0.563)**	**(0.563)**	**(0.540)**	**(0.564)**	(0.774)	(0.778)
Complementarity * Department Quality	**0.129****	**0.130****	**0.196*****	**0.196*****	-0.170	-0.169	**0.331****	**0.330****
	(0.058)	**(0.058)**	**(0.063)**	**(0.063)**	(0.133)	(0.133)	**(0.140)**	**(0.139)**
Linear Time Trend		0.073		-0.073		0.610		0.175
		(0.122)		(0.127)		(0.406)		(0.271)
Post '05	0.654	0.435	0.452	0.672	**1.798***	-0.022	-0.796	-1.286
	(0.513)	(0.696)	(0.636)	(0.831)	**(1.036)**	(1.795)	(0.547)	(1.046)
Post '05 * Department Quality	**-1.595****	**-1.598****	-1.580	-1.578	**-1.266***	**-1.282***	-1.089	-1.088
	(0.720)	**(0.720)**	(0.998)	(0.998)	**(0.745)**	**(0.744)**	(0.675)	(0.674)
Post '05 * Complementarity	-0.270	-0.271	-0.077	-0.077	**-1.063***	**-1.055***	**0.628****	**0.634****
	(0.258)	(0.258)	(0.299)	(0.299)	**(0.584)**	**(0.585)**	**(0.313)**	**(0.315)**
Post '05 * Complementarity * Department Quality	**0.442*****	**0.442*****	**0.447****	**0.447****	**0.387***	**0.390***	0.086	0.085
	(0.165)	**(0.165)**	**(0.225)**	**(0.225)**	**(0.200)**	**(0.200)**	(0.166)	(0.166)
N	2673	2673	2280	2280	1347	1347	1758	1758

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. In columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Post is 0 before 2005 and 1 thereafter. All regressions contain field fixed effects. Standard errors are robust, clustered by department.

Table 5.4: Pre-trend Test I

(Dep.Var.: Average Productivity)	All new hires		Junior hires		Senior Hires		Leavers	
	1_a	1_b	2_a	2_b	3_a	3_b	4_a	4_b
Department Quality	-0.031	-0.031	-0.485	-0.486	**1.264****	**1.267****	-0.050	-0.018
	(0.264)	(0.264)	(0.304)	(0.304)	**(0.631)**	**(0.627)**	(0.674)	(0.671)
Complementarity	**1.384*****	**1.386*****	**1.318****	**1.317****	**1.642*****	**1.670*****	0.867	0.926
	(0.447)	**(0.447)**	**(0.568)**	**(0.571)**	**(0.605)**	**(0.605)**	(0.787)	(0.788)
Complementarity* Department Quality	0.084	0.084	**0.179****	**0.179****	-0.223	-0.224	0.251	0.244
	(0.062)	(0.062)	**(0.071)**	**(0.071)**	(0.151)	(0.150)	(0.168)	(0.167)
Linear Time Trend		0.065		0.044		0.228		0.598
		(0.216)		(0.272)		(0.451)		(0.413)
post '04	0.150	0.021	0.074	-0.014	0.422	-0.010	1.184	0.044
	(0.562)	(0.726)	(0.307)	(0.655)	(1.977)	(2.307)	(1.556)	(1.765)
Post '05	0.564	0.465	0.422	0.355	1.433	1.086	-1.554	**-2.421***
	(0.680)	(0.801)	(0.663)	(0.851)	(1.872)	(2.007)	(1.252)	**(1.421)**
Post '04 * Department Quality	-0.548	-0.548	-0.197	-0.196	-0.462	-0.471	-0.495	-0.533
	(0.350)	(0.351)	(0.276)	(0.276)	(1.077)	(1.075)	(0.882)	(0.882)
Post '05 * Department Quality	-1.221	-1.224	-1.459	-1.461	-0.910	-0.911	-0.793	-0.779
	(0.775)	(0.775)	(1.012)	(1.010)	(0.991)	(0.991)	(0.836)	(0.837)
Post '04 * Complementarity	-0.049	-0.049	-0.067	-0.067	-0.279	-0.283	-0.934	-0.943
	(0.300)	(0.300)	(0.166)	(0.167)	(1.115)	(1.114)	(0.988)	(0.987)
Post '05 * Complementarity	-0.239	-0.240	-0.037	-0.038	-0.826	-0.825	1.245	1.258
	(0.342)	(0.342)	(0.313)	(0.312)	(1.079)	(1.080)	(0.801)	(0.801)
Post '04 * Complementarity * Department Quality	0.134	0.134	0.031	0.031	0.249	0.251	0.225	0.233
	(0.100)	(0.100)	(0.065)	(0.065)	(0.331)	(0.331)	(0.285)	(0.285)
Post '05 * Complementarity * Department Quality	**0.349***	**0.350***	**0.430***	**0.430***	0.195	0.196	-0.059	-0.063
	(0.185)	**(0.185)**	**(0.231)**	**(0.230)**	(0.314)	(0.314)	(0.251)	(0.251)
N	2673	2673	2280	2280	1347	1347	1758	1758

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. In columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Post'05 is 0 before 2005 and 1 thereafter. Post'04 is 0 before 2004 and 1 thereafter. All regressions contain field fixed effects. Standard errors are robust, clustered by department.

Table 5.5: Pre-trend Test II

(Dep.Var.: Average Productivity)	All new hires		Junior hires		Senior Hires		Leavers	
	1_a	1_b	2_a	2_b	3_a	3_b	4_a	4_b
Department Quality	0.676	0.676	0.030	0.029	1.889**	1.884**	-1.558	-1.573
	(0.452)	(0.452)	(0.454)	(0.456)	(0.924)	(0.925)	(1.616)	(1.621)
Complementarity (pre-sample, 96-00)	1.626***	1.627***	1.471**	1.472**	2.245***	2.260***	4.249	4.251
	(0.534)	(0.534)	(0.667)	(0.667)	(0.831)	(0.828)	(2.724)	(2.723)
Complementarity * Department Quality	-0.075	-0.075	0.078	0.078	-0.423**	-0.422**	0.437	0.438
	(0.104)	(0.104)	(0.110)	(0.110)	(0.204)	(0.205)	(0.373)	(0.374)
Linear Time Trend		0.077		-0.140		0.211		1.967*
		(0.593)		(0.754)		(0.945)		(1.082)
post '02	0.828	0.751	0.371	0.510	1.419	1.208	-0.220	-2.196
	(0.611)	(0.880)	(0.661)	(1.006)	(1.533)	(1.813)	(1.487)	(2.000)
Post '03	-0.790	-0.868	-0.914	-0.774	0.746	0.535	0.019	-1.942
	(0.579)	(0.760)	(0.565)	(0.849)	(1.757)	(1.920)	(1.315)	(1.694)
post '04	0.413	0.335	0.572	0.711	-0.579	-0.791	1.496	-0.471
	(0.701)	(0.978)	(0.490)	(0.965)	(2.552)	(2.837)	(1.916)	(2.332)
Post '05	0.568	0.451	0.441	0.654	1.426	1.105	-1.680	-4.527**
	(0.681)	(1.227)	(0.664)	(1.463)	(1.878)	(2.467)	(1.463)	(1.812)
Post '02 * Department Quality	-1.118*	-1.118*	-1.198**	-1.198**	-0.391	-0.390	2.924*	2.918*
	(0.596)	(0.596)	(0.594)	(0.594)	(1.191)	(1.191)	(1.755)	(1.756)
Post '03 * Department Quality	0.113	0.113	0.883	0.884	-1.476	-1.478	-0.928	-0.927
	(0.617)	(0.617)	(0.585)	(0.585)	(1.568)	(1.568)	(1.591)	(1.591)
Post '04 * Department Quality	-0.262	-0.262	-0.427	-0.427	0.730	0.731	-0.785	-0.782
	(0.507)	(0.507)	(0.422)	(0.422)	(1.743)	(1.743)	(1.577)	(1.578)
Post '05 * Department Quality	-1.229	-1.232	-1.486	-1.480	-0.899	-0.900	-1.463	-1.492
	(0.776)	(0.777)	(1.014)	(1.010)	(0.995)	(0.995)	(1.288)	(1.292)
Post '02 * Complementarity	-0.466	-0.466	-0.250	-0.251	-0.739	-0.738	-0.457	-0.457
	(0.319)	(0.319)	(0.349)	(0.349)	(0.745)	(0.746)	(0.750)	(0.750)
Post '03 * Complementarity	0.388	0.388	0.383	0.383	-0.295	-0.296	0.106	0.104
	(0.297)	(0.297)	(0.279)	(0.279)	(0.921)	(0.922)	(0.731)	(0.732)
Post '04 * Complementarity	-0.165	-0.165	-0.254	-0.254	0.179	0.179	-1.139	-1.139
	(0.379)	(0.380)	(0.264)	(0.264)	(1.399)	(1.399)	(1.209)	(1.210)
Post '05 * Complementarity	-0.240	-0.241	-0.042	-0.039	-0.822	-0.821	1.446	1.461
	(0.342)	(0.341)	(0.314)	(0.311)	(1.082)	(1.083)	(0.917)	(0.916)
Post '02 * Complementarity * Department Quality	0.229*	0.229*	0.209	0.209	0.170	0.170	-0.469	-0.468
	(0.130)	(0.130)	(0.137)	(0.137)	(0.255)	(0.255)	(0.396)	(0.396)
Post '03 * Complementarity * Department Quality	0.019	0.019	-0.121	-0.121	0.339	0.340	0.180	0.180
	(0.145)	(0.145)	(0.135)	(0.135)	(0.383)	(0.383)	(0.349)	(0.349)
Post '04 * Complementarity * Department Quality	0.048	0.048	0.049	0.049	-0.050	-0.050	0.280	0.279
	(0.135)	(0.135)	(0.099)	(0.099)	(0.484)	(0.484)	(0.410)	(0.411)
Post '05 * Complementarity * Department Quality	0.351*	0.351*	0.436*	0.434*	0.193	0.193	0.067	0.072
	(0.186)	(0.186)	(0.231)	(0.230)	(0.315)	(0.315)	(0.337)	(0.338)
N	2673	2673	2280	2280	1347	1347	1758	1758

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. In columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Post'0# is 0 before 200# and 1 thereafter. All regressions contain field fixed effects. Standard errors are robust, clustered by department.

Table 5.6: Controlling for Hiring Budget

(Dep.Var.: Average Productivity)	All new hires		Junior hires		Senior Hires		Leavers	
	1_a	1_b	2_a	2_b	3_a	3_b	4_a	4_b
Department Quality	-0.243	-0.246	**-0.589****	**-0.586****	**1.128****	**1.118***	-0.319	-0.318
	(0.239)	(0.240)	**(0.256)**	**(0.256)**	**(0.574)**	**(0.573)**	(0.538)	(0.538)
Complementarity	**1.352*****	**1.348*****	**1.261****	**1.269****	**1.511*****	**1.555*****	0.591	0.591
	(0.437)	**(0.440)**	**(0.562)**	**(0.563)**	**(0.561)**	**(0.590)**	(0.769)	(0.772)
Complementarity * Department Quality	**0.134****	**0.135****	**0.199*****	**0.198*****	-0.162	-0.158	**0.349****	**0.349****
	(0.059)	**(0.059)**	**(0.064)**	**(0.064)**	(0.136)	(0.136)	**(0.141)**	**(0.141)**
Hiring Budget	-0.024	-0.025	-0.013	-0.012	-0.044	-0.051	-0.056	-0.058
	(0.016)	(0.016)	(0.020)	(0.021)	(0.046)	(0.046)	(0.037)	(0.037)
Linear Time Trend		0.086		-0.065		0.634		0.199
		(0.124)		(0.129)		(0.408)		(0.276)
Post '05	-0.588	-0.857	-0.306	-0.098	-0.417	-2.398	**-1.998****	**-2.579***
	(0.774)	(0.962)	(0.977)	(1.191)	(1.510)	(2.142)	**(0.930)**	**(1.355)**
Post '05 * Department Quality	**-1.472****	**-1.474****	-1.501	-1.500	-1.085	-1.094	-0.977	-0.973
	(0.716)	**(0.716)**	(0.993)	(0.993)	(0.760)	(0.760)	(0.676)	(0.675)
Post '05 * Complementarity	-0.183	-0.182	-0.024	-0.025	-0.915	-0.901	**0.691****	**0.699****
	(0.264)	(0.265)	(0.312)	(0.312)	(0.568)	(0.570)	**(0.318)**	**(0.320)**
Post '05 * Complementarity * Department Quality	**0.408****	**0.408****	**0.426***	**0.426***	**0.338***	**0.339***	0.055	0.054
	(0.164)	**(0.164)**	**(0.225)**	**(0.225)**	**(0.197)**	**(0.197)**	(0.167)	(0.167)
Post '05 * Hiring Budget	**0.101****	**0.102****	0.061	0.060	**0.164***	**0.171***	**0.101***	**0.103***
	(0.041)	**(0.042)**	(0.053)	(0.054)	**(0.096)**	**(0.096)**	**(0.057)**	**(0.057)**
N	2673	2673	2280	2280	1347	1347	1758	1758

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006. In columns 1a and 1b the dependent variable is the average productivity of all new hires of department j in year t, in columns 2a and 2b the average productivity of junior hires, in columns 3a and 3b the average productivity of senior hires and in columns 4a and 4b the average productivity of leavers. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Budget is the number of affiliates of a university turning 66 in year t. Post is 0 before 2005 and 1 thereafter. All regressions contain field fixed effects. Standard errors are robust, clustered by department.

Table 5.7: Robustness Checks

(Dep.Var.: Average Productivity)	**FE model**		**RE w/ year FE**	**RE 2000-2009**		**RE w/ Dep. Budget**	
	1_a	**1_b**	**2_a**	**3_a**	**3_b**	**4_a**	**4_b**
Department Quality			-0.236	-0.270	-0.272	-0.236	-0.238
			(0.239)	(0.219)	(0.220)	(0.239)	(0.240)
Complementarity			**1.389*** **	**1.101*** **	**1.097*** **	**1.360*** **	**1.356*** **
			(0.438)	**(0.315)**	**(0.316)**	**(0.438)**	**(0.441)**
Complementarity * Department Quality			**0.131** **	**0.139*** **	**0.140*** **	**0.132** **	**0.133** **
			(0.058)	**(0.053)**	**(0.053)**	**(0.059)**	**(0.059)**
Departmental Hiring Budget						-0.046	-0.048
						(0.087)	(0.088)
Linear Time Trend		0.101			0.071		0.077
		(0.127)			(0.086)		(0.123)
Post '05	0.650	0.351		0.445	0.112	0.135	-0.099
	(0.588)	(0.787)		(0.348)	(0.509)	(0.638)	(0.818)
Post '05 * Department Quality	**-1.600** **	**-1.605** **	**-1.597** **	**-0.949** **	**-0.951** **	**-1.537** **	**-1.539** **
	(0.727)	**(0.726)**	**(0.721)**	**(0.428)**	**(0.429)**	**(0.720)**	**(0.720)**
Post '05 * Complementarity	-0.213	-0.215	-0.268	-0.140	-0.138	-0.239	-0.239
	(0.299)	(0.298)	(0.255)	(0.175)	(0.175)	(0.262)	(0.262)
Post '05 * Complementarity * Department Quality	**0.437*** **	**0.438*** **	**0.443*** **	**0.281*** **	**0.282*** **	**0.421** **	**0.421** **
	(0.167)	**(0.167)**	**(0.165)**	**(0.101)**	**(0.101)**	**(0.165)**	**(0.165)**
Post '05 * Departmental Hiring Budget						**0.366***	**0.369***
						(0.219)	**(0.219)**
N	2673	2673	2673	4240	4240	2673	2673

Notes: * denotes significance at 10%, ** at 5% and *** at 1%. The unit of observation is department j. The sample includes all departments of German public universities in the years 2001 to 2006, except in columns 3a and 3b where the sample spans 2001-2009. In all columns the dependent variable is the average productivity of all new hires of department j in year t. These average productivities are calculated as the average impact-factor weighted number of publications in t-6 to t-1. Department quality is measured as the pre-sample mean average productivity of a department's affiliates. Complementarity is measured as the average number of authors on a paper in a field (calculated over the pre-sample years 1996-2000). Departmental budget is the number of affiliates of a department turning 66 in year t. Post is 0 before 2005 and 1 thereafter. Columns 1a and 1b show the results of the baseline specification estimated as FE model, Column 2 estimates the baseline specification as RE model with year fixed effects, Columns 3a and 3b estimates the baseline model with a balanced sample (2001-2009) and in Columns 4a and 4b a proxy for the hiring budget at the departmental level is added to the baseline model. All regressions except 1a and 1b contain field fixed effects. Standard errors are robust, clustered by department.

Appendix A: Proofs

Proof of Proposition 1:

Sort i by productivity, so that $\theta_i \geq \theta_{i+1}$. For ease of notation, I will refer to an academic with his productivity type rank order number, so that $\theta_{i=1} > \theta_j, \forall j \neq i$, $\theta_{i=m} < \theta_j, \forall j \neq i$ and $k > l$ iff $\theta_k < \theta_l$. The unique stable matching $\tilde{\mu}$ of the baseline model is then given by:

$$\tilde{\mu} = \begin{cases} \tilde{\mu}(1) = 2, \tilde{\mu}(3) = 4, .., \tilde{\mu}(k+2) = k+1, \tilde{\mu}(k) = k, .., \tilde{\mu}(m) = m & \text{if } k \text{ odd} \\ \tilde{\mu}(1) = 2, \tilde{\mu}(3) = 4, .., \tilde{\mu}(k+1) = k+1, \tilde{\mu}(k) = k, .., \tilde{\mu}(m) = m & \text{if } k \text{ even} \end{cases} \tag{5.5.1}$$

where:

$$k = \begin{cases} i : u_{k|k+1} < \underline{u}, u_{k+1|k} \geq \underline{u} & \text{for } i \text{ even} \\ i : u_{k|k-1} < \underline{u}, u_{k+1|k+2} \geq \underline{u} & \text{for } i \text{ odd} \end{cases} \tag{5.5.2}$$

It is immediate that no blocking pair exists for this matching, but that at least one blocking pair exists for any other matching. It is also immediate that the matching $\tilde{\mu}$ maximises total surplus and that this is a unique maximum if $f(\theta_i \mid \theta_j)$ exhibits *increasing differences*. QED

Proof of Proposition 2:

I first show that the matching $\tilde{\mu}$ that was stable in the baseline is now not stable with positive probability. To see this, keep i sorted by productivity, so that $\theta_i \geq \theta_{i+1}$ as in the baseline case. Define $D^{\theta}_{k(i,j)} \equiv f(\theta_k \mid \theta_i) - f(\theta_k \mid \theta_j)$ and $D^{\nu}_{k(i,j)} = \nu_{k,j} - \nu_{k,i}$. Then i and $i+t$ form a blocking pair if $\alpha f\left(\theta_i \mid \theta_{\tilde{\mu}(i)}\right) + \nu_{i,\tilde{\mu}(i)} \leq \alpha f(\theta_i \mid \theta_{i+1}) + \nu_{i,i+1}$ and $\alpha f(\theta_{i+t} \mid \theta_i) + \nu_{i+t,i} \geq \alpha f\left(\theta_{i+t} \mid \theta_{\tilde{\mu}(i+t)}\right) + \nu_{i+t,\tilde{\mu}(i+t)}$. Here, $\tilde{\mu}(i) = i+1$ if i is odd, and $\tilde{\mu}(i) = i-1$ if i is even. Similarly, $\tilde{\mu}(i+t) = i+t+1$ if $i+t$ is odd, and $\tilde{\mu}(i) = i+t-1$ if $i+t$ is even. So we have that $\tilde{\mu}$ is not stable if $\exists i, i+t$ such that:

$$\alpha D^{\theta}_{i(\tilde{\mu}(i),i+t)} \leq D^{\nu}_{i(i+t,\tilde{\mu}(i))} \tag{5.5.3}$$

and

$$\alpha D^{\theta}_{i+t(i,\tilde{\mu}(i+t))} \geq D^{\nu}_{i+t(\tilde{\mu}(i+t),i)} \tag{5.5.4}$$

where $t > 1$ if i is odd and $t > 0$ if i is even, and at least one inequality strict. If assumption 2 holds, these equations can be rewritten as:

$$\alpha c_2\left(\theta_{\tilde{\mu}(i)}-\theta_{i+t}\right)\leq D^{\nu}_{i(i+t,\tilde{\mu}(i))} \tag{5.5.5}$$

and

$$\alpha c_2\left(\theta_i-\theta_{\tilde{\mu(i+t)}}\right)\geq D^{\nu}_{i+t(\tilde{\mu(i+t)},i)} \tag{5.5.6}$$

The probability that 5.5.5 holds is given by $P_i = Pr\left(\alpha c_2\left(\theta_{\tilde{\mu}(i)}-\theta_{i+t}\right)\leq D^{\nu}_{i(i+t,\tilde{\mu}(i))}\right)=1-Pr\left(\alpha c_2\left(\theta_{\tilde{\mu}(i)}-\theta_{i+t}\right)\right)=1-Pr\left(\alpha D^{\theta}_{i(\tilde{\mu}(i),i+t)}\right)$, while the probability that 5.5.6 holds is given by $P_{i+t}=Pr\left(\alpha D^{\theta}_{i+t(i,\tilde{\mu(i+t)})}\geq D^{\nu}_{i+t(\tilde{\mu(i+t)},i)}\right)=Pr\left(\alpha D^{\theta}_{i+t(i,\tilde{\mu(i+t)})}\right)=$ $Pr\left(\alpha c_2\left(\theta_i-\tilde{\mu(i+t)}\right)\right)$. Here, $Pr(.)=H_{\nu_{jk}-\nu_{jl}}(.)$ is the cdf of $\nu_{jk}-\nu_{jl}$, the difference between the noise draws when j is matched with k and when j is matched with l, and $\nu_{jk}-\nu_{jl}\sim h\left(\nu_{jk}-\nu_{jl}\right)$. For ease of notation, let $x\equiv\theta_{\tilde{\mu}(i)}-\theta_{i+t}$ and $y\equiv\theta_i-\theta_{\tilde{\mu(i+t)}}$. The probability that $i, i+t$ is a blocking pair to baseline stable matching $\tilde{\mu}$ is then given by:

$$P_{i,i+t}=\left(1-H\left(\alpha c_2 x\right)\right)H\left(\alpha c_2 y\right) \tag{5.5.7}$$

and the probability that $\tilde{\mu}$ is not stable is

$$P\left(\tilde{\mu}\ unstable\right)=\sum_{i,i+t}P_{i,i+t} \tag{5.5.8}$$

. By assumptions 1 and 2, both terms in 5.5.7 are positive, and hence both 5.5.7 and 5.5.8 are positive. Therefore, the baseline stable matching $\tilde{\mu}$ is no longer stable with non-zero probability.

I next show that the probability that the baseline stable matching $\tilde{\mu}$ is not stable is smaller if complementarities are stronger. Recall that $f(.\mid .)$ exhibits stronger complementarities than $\tilde{f}(.\mid .)$ if $\left\|D^{\theta}_{k(i,j)}\right\|=\|f(\theta_k\mid\theta_i)-f(\theta_k\mid\theta_j)\|>\left\|\tilde{f}(\theta_k\mid\theta_i)-\tilde{f}(\theta_k\mid\theta_j)\right\|$ $=\left\|\tilde{D}^{\theta}_{k(i,j)}\right\|$ for $i>j$. Under assumption 1, this means that $f_2=c_2>\tilde{c_2}=\tilde{f}_2$. The probability that $\tilde{\mu}$ is not stable is smaller if complementarities are stronger, if $\frac{\partial P(\tilde{\mu}\ unstable)}{\partial c_2}<0$. We have that

$$\frac{\partial P_{i,i+t}}{\partial c_2}=-h\left(\alpha c_2 x\right)H\left(\alpha c_2 y\right)\alpha x+\left(1-H\left(\alpha c_2 x\right)\right)h\left(\alpha c_2 y\right)\alpha y \tag{5.5.9}$$

Axiom 1: $\forall i, i+t, \exists j, j+t$ with $i<j$ such that $x_{i,i+t}=\theta_{\tilde{\mu}(i)}-\theta_{i+t}=\theta_j-\theta_{\tilde{\mu(j+t)}}=y_{j,j+t}$ and $y_{i,i+t}=\theta_i-\theta_{\tilde{\mu(i+t)}}=\theta_{\tilde{\mu}(j)}-\theta_{j+t}=x_{j,j+t}$

For any $i, i+t$ and $j, j+t$ with $i < j$ and $x_{i,i+t} = y_{j,j+t}$ and $y_{i,i+t} = x_{j,j+t}$ we can write the probability that $i, i+t$ or $j, j+t$ is a blocking pair to $\tilde{\mu}$ as:

$$P_{i,i+t} + P_{j,j+t} = \left(1 - H\left(\alpha c_2 x_{i,i+t}\right)\right) H\left(\alpha c_2 y_{i,i+t}\right) + \left(1 - H\left(\alpha c_2 y_{i,i+t}\right)\right) H\left(\alpha c_2 x_{i,i+t}\right) \tag{5.5.10}$$

The change of this probability with respect to the complementarity constant c_2 is then:

$$\begin{aligned} \tfrac{\partial P_{i,i+t}}{\partial c_2} + \tfrac{\partial P_{j,j+t}}{\partial c_2} &= h\left(\alpha c_2 x\right) \alpha x \left[\left(1 - H\left(\alpha c_2 y\right)\right) - H\left(\alpha c_2 y\right)\right] \\ &+ h\left(\alpha c_2 y\right) \alpha y \left[\left(1 - H\left(\alpha c_2 x\right)\right) - H\left(\alpha c_2 x\right)\right] \end{aligned} \tag{5.5.11}$$

By assumption 2, all arguments of $h()$ and $H()$ and are in the support of $h()$ and therefore $h\left(\alpha c_2 x\right), H\left(\alpha c_2 x\right) > 0$, $h\left(\alpha c_2 y\right), H\left(\alpha c_2 y\right) > 0$ and $H\left(\alpha c_2 x\right) > \left(1 - H\left(\alpha c_2 x\right)\right)$ because $x, y > 0$ and $h()$ is symmetric around zero. Therefore, the two terms in square brackets in 5.5.11 are negative and hence so is 5.5.11. By *Axiom* 1 it then follows that

$$\frac{\partial \sum_{i,i+t} P_{i,i+t}}{\partial c_2} < 0 \tag{5.5.12}$$

and hence that the probability that $\tilde{\mu}$ is not stable is smaller when complementarities are stronger.

As shown by Gale and Shapley (1962), one-sided matching models without transferable utility might not have a stable matching. It is straightforward to come up with examples of noise draws for which there is no stable matching for the academic job market either. If a stable matching exists and if the maximal positive assortative matching is not stable, it is immediate that a less than maximal positive assortative matching is stable, since the maximal positive assortative matching is unique. If no stable matching exists, and under the blocking dynamics introduced in Roth and Vate (1990) adapted to roommate markets as in Diamantoudi et al. (2004), it follows from the definition of the blocking dynamics and absorbing sets that if no stable matching exists, the absorbing set is not a singleton (Klaus et al. 2010). Given that the maximal positive assortative matching is unique, the absorbing set must contain at least one less than maximally positive assortative matching if no stable matching exists. Through the blocking dynamics such a less than maximally positive assortative matching is reached with positive probability. QED

Proof of Proposition 3:

I first show that, if α is larger post-reform, so that $\alpha^{post} > \alpha^{pre}$, blocking pairs to the maximal positive assortative matching $\tilde{\mu}$ exist with a smaller probability than in the pre-reform scenario. The proof follows the same reasoning as the proof for the effect of stronger complementarities on stability of baseline stable matching $\tilde{\mu}$ in the proof of proposition 2 above. The probability that $\tilde{\mu}$ is not stable is smaller if α is larger post-reform; if $\frac{\partial P(\tilde{\mu} unstable)}{\partial \alpha} < 0$. We have that

$$\frac{\partial P_{i,i+t}}{\partial \alpha} = -h\left(\alpha c_2 x\right) H\left(\alpha c_2 y\right) c_2 x + \left(1 - H\left(\alpha c_2 x\right)\right) h\left(\alpha c_2 y\right) c_2 y \tag{5.5.13}$$

For any $i, i+t$ and $j, j+t$ with $i < j$ and $x_{i,i+t} = y_{j,j+t}$ and $y_{i,i+t} = x_{j,j+t}$ we can write the probability that $i, i+t$ or $j, j+t$ is a blocking pair to $\tilde{\mu}$ as:

$$P_{i,i+t} + P_{j,j+t} = \left(1 - H\left(\alpha c_2 x_{i,i+t}\right)\right) H\left(\alpha c_2 y_{i,i+t}\right) + \left(1 - H\left(\alpha c_2 y_{i,i+t}\right)\right) H\left(\alpha c_2 x_{i,i+t}\right) \tag{5.5.14}$$

The change of this probability with respect to the proportionality constant α is then:

$$\begin{aligned} \tfrac{\partial P_{i,i+t}}{\partial \alpha} + \tfrac{\partial P_{j,j+t}}{\partial \alpha} &= h\left(\alpha c_2 x\right) c_2 x \left[\left(1 - H\left(\alpha c_2 y\right)\right) - H\left(\alpha c_2 y\right)\right] \\ &+ h\left(\alpha c_2 y\right) c_2 y \left[\left(1 - H\left(\alpha c_2 y\right)\right) - H\left(\alpha c_2 y\right)\right] \end{aligned} \tag{5.5.15}$$

By assumption 2, all arguments of $h()$ and $H()$ are in the support of $h()$ and therefore $h\left(\alpha c_2 x\right), H\left(\alpha c_2 x\right)$,
$h\left(\alpha c_2 y\right), H\left(\alpha c_2 y\right) > 0$ and $H\left(\alpha c_2 x\right) > \left(1 - H\left(\alpha c_2 x\right)\right)$ because $x, y > 0$ and $h()$ is symmetric around zero. Therefore, the two terms in square brackets in 5.5.15 are negative and hence so is 5.5.15. By *Axiom* 1 it then follows that

$$\frac{\partial \sum_{i,i+t} P_{i,i+t}}{\partial \alpha} < 0 \tag{5.5.16}$$

and hence the probability that $\tilde{\mu}$ is not stable is smaller when the proportionality constant α is larger.

In order to show that the decrease in the probability that $\tilde{\mu}$ is not stable with respect to the proportionality constant α is larger if complementarities are stronger, I derive the cross-derivative of this probability with respect to α and c_2 and show that it is negative.

This cross-derivative is given by:

$$\begin{aligned}\frac{\partial^2 P_{i,i+t}}{\partial\alpha\partial c_2} = &-\left[xh\left(\alpha c_2 x\right)+\alpha c_2 x^2 h'\left(\alpha c_2 x\right)\right] H\left(\alpha c_2 y\right) \\ &+\left[xh\left(\alpha c_2 y\right)+\alpha c_2 y^2 h'\left(\alpha c_2 y\right)\right]\left(1-H\left(\alpha c_2 x\right)\right) \\ &-2\alpha c_2 xyh\left(\alpha c_2 x\right) h\left(\alpha c_2 y\right)\end{aligned} \tag{5.5.17}$$

It is straightforward, though cumbersome, to show that this cross-derivative is negative for e.g. a uniform distribution of the idiosyncratic noise terms: $\nu_{ij} \sim U\left[-\nu,\nu\right]$, with $\nu = \alpha c_2\left(\bar{\theta}-\underline{\theta}\right)$.

Secondly, I need to show that any matching μ that was stable pre-reform and that matches academics with a productivity rank difference greater than one is less likely to be stable post-reform. To see this, suppose that pre-reform a matching $\hat{\mu}$ was stable in which i was matched with $i+2$, and $i+1$ with $i+3$ while all other pairings in $\hat{\mu}$ were as in $\tilde{\mu}$. A blocking pair i,j with $j \neq i+2$ to the matching $\hat{\mu}$ exists if:

$$\alpha D^{\theta}_{i(i+2,j)} \leq D^{\nu}_{i(j,i+2)} \tag{5.5.18}$$

and

$$\alpha D^{\theta}_{j(i,\hat{\mu}(j))} \geq D^{\nu}_{j(\hat{\mu}(j),i)} \tag{5.5.19}$$

Suppose $j=i+1$, so $\hat{\mu}(j)=i+3$. We then have that $D^{\theta}_{i(i+2,j)} = D^{\theta}_{i(i+2,i+1)} < 0$ and the post-reform constraint on the higher productivity academic of the potential blocking pair (inequality 5.5.18) is relaxed if $\alpha_i^{post} > \alpha_i^{pre}$. Hence the probability that i prefers being matched to $\tilde{\mu}\left(i\right)$ instead of $\hat{\mu}(i)$ is larger post-reform. At the same time, $D^{\theta}_{j(i,\hat{\mu}(j))} = D^{\theta}_{i+1(i,i+3)} > 0$, so that the post-reform constraint on the lower productivity academic of the potential blocking pair (inequality 5.5.19) is relaxed too, so the probability that j prefers being matched to $\tilde{\mu}\left(j\right)$ instead of $\hat{\mu}(j)$ is larger post-reform. It is thus more likely that $i,i+1$ is a blocking pair for $\hat{\mu}$ post-reform. Coincidentally, if $\{i,i+1\}$ is a blocking pair to $\hat{\mu}$, then $\{i,i+2\}$ is not a blocking pair to $\tilde{\mu}$. It can be shown that the same holds for any $j \neq i+2$, and for any non-maximal matching $\hat{\mu}$ in which the matching partner of an academic i in the maximal positive assortative matching $\tilde{\mu}$ (i.e.$i \pm 1$) is swapped for the next-closest in rank academic ($i \pm 2$) and i's matching partner in $\tilde{\mu}$ is matched to the matching partner in $\tilde{\mu}$ of i's matching partner in $\hat{\mu}$ (and any iteration of this swap). By the same reasoning, it follows that the probability that $i,i+1$ is a blocking pair for $\hat{\mu}$

post-reform is larger if complementarities are stronger. QED

6 Conclusion and Discussion

This thesis studies the effort and selection effect of performance pay in academia, as well as the effect of performance pay on matching assortativeness. In order to study these effects, I use the introduction of performance pay in German academia as a natural experiment. I constructed a new panel data set that encompasses the universe of German academics and contains information on their affiliations and research productivity for the purposes of this study. This data set spans a 15-year period, from 1999 to 2013, and, importantly, includes both pre-reform and post-reform years.

I find that performance pay can have a significant and sizeable positive effect on effort in academia, estimating a 35% increase in productivity. Because academics are thought to be intrinsically motivated (McCormack et al. 2014, Besley and Ghatak 2005) and because extrinsic incentives might crowd out intrinsic motivation (Ariely et al. 2009, Dickinson and Villeval 2008, Frey and Jegen 2001), this is not a trivial or ex ante obvious finding.

I also provide evidence of a positive selection effect. I find that higher productivity academics are more likely to select into performance pay, and that this effect is stronger for younger academics. This result is reassuring, because the extrinsic motivation introduced by performance pay may crowd out intrinsic motivation with respect to agents' sorting decision (e.g. Georgellis et al. 2010). If less motivated or lower productivity academics were attracted to performance pay, performance pay might have an overall negative effect on academic output, despite the positive effort effect. Fortunately, I find evidence of a positive selection effect, in line with the prediction of Besley and Ghatak (2006).

Furthermore, this thesis provides evidence that performance pay increases positive assortative matching by productivity in academia two- to threefold. If there are increasing differences in production in academia, this increase in positive assortative matching increases total academic output (Legros and Newman 2002). On the other hand, this increased clustering of highly productive academics might imply a less equal distribution of educational quality and local knowledge spillovers through university-industry partnerships. It would therefore be interesting to study how the change in matching assortativeness in academia affects outcomes in tertiary education. Similarly, it would be interesting to assess to what extent the increased clustering of highly productive academics affects the distribution of localised spillovers of academic research and, ultimately, how this affects technological progress (cf. Griliches 1998).

Given that I find both a positive effort effect and a positive selection effect, performance pay appears to be an effective instrument for increasing academic research output. The question remains however, how useful this research is for technological progress and growth, especially in the longer run. The performance pay scheme introduced in Germany links pay to i.a. weighted publication counts (Detmer and Preissler 2004 and 2005, BMBF 2002). It could be that academics shift attention to research projects that are less risky (more likely to yield a publication) and shorter-term, so as to increase the number of publications and assure a more or less steady stream of publications, and thereby increase the chances of earning performance bonuses. This might imply a move away from basic towards more applied research and from more unconventional, possibly ground-breaking research towards more incremental research. To the extent that we need such basic and groundbreaking research for technological progress, performance pay with explicit, output-based performance criteria as studied here might, though increasing academic output, ultimately slow down growth (Bresnahan and Trajtenberg 1995, Griliches 1958). It would therefore be interesting to study if and how research output changes under different kinds of performance pay schemes.

Another aspect of performance pay in academia that would be worth exploring is how it impacts collaboration. Tournament schemes might, for instance, have a very different effect on co-authorship than piece-rate schemes. Under a tournament scheme, academics might be less willing to collaborate with academics whom they compete with (cf. Harbring and Irlenbusch 2011 and 2008, Chen 2003) and they might consequently seek co-authors at different universities, in different states or even countries, or prefer to co-author with academics from a different cohort or hierarchy level. This could have interesting consequences for the co-authorship network and, by extension, the diffusion of (academic) knowledge.

Bibliography

Acemoglu, D., Aghion, P. and Zilibotti, F.: 2006, Distance to frontier, selection, and economic growth, *Journal of the European Economic association* **4**(1), 37–74.

Adams, J.: 1990, Fundamental stocks of knowledge and productivity growth, *Journal of Political Economy* **98**(4), 673–702.

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

Agrawal, A.: 2001, University-to-industry knowledge transfer: literature review and unanswered questions, *International Journal of management reviews* **3**, 285–302.

Agrawal, A., McHale, J. and Oettl, A.: 2014, Why stars matter, *Technical report*, National Bureau of Economic Research.

Angrist, J., Lang, D. and Oreopoulos, P.: 2009, Incentives and services for college achievement: Evidence from a randomized trial, *American Economic Journal: Applied Economics* pp. 136–163.

Angrist, J. and Lavy, V.: 2009, The effects of high stakes high school achievement awards: Evidence from a randomized trial, *The American Economic Review* pp. 1384–1414.

Ariely, D., Bracha, A. and Meier, S.: 2009, Doing good or doing well? image motivation and monetary incentives in behaving prosocially, *The American Economic Review* pp. 544–555.

Audretsch, D. B. and Feldman, M. P.: 1996, R&d spillovers and the geography of innovation and production, *The American economic review* pp. 630–640.

Audretsch, D. B. and Feldman, M. P.: 2004, Knowledge spillovers and the geography of innovation, *Handbook of regional and urban economics* **4**, 2713–2739.

Azoulay, P., Zivin, J. S. G. and Wang, J.: 2010, Superstar extinction, *The Quarterly Journal of Economics* **125**(2), 549–589.

Bandiera, O., Barankay, I. and Rasul, I.: 2005, Social preferences and the response to incentives: Evidence from personnel data, *The Quarterly Journal of Economics* pp. 917–962.

Bandiera, O., Barankay, I. and Rasul, I.: 2013, Team incentives: Evidence from a firm level experiment, *Journal of the European Economic Association* **11**(5), 1079–1114.

Becker, G. S.: 1973, A theory of marriage, part i, *The Journal of Political Economy* pp. 813–846.

Belenzon, S. and Schankerman, M.: 2009, University knowledge transfer: private ownership, incentives, and local development objectives, *Journal of Law and Economics* **52**(1), 111–144.

Bellemare, C., Lepage, P. and Shearer, B.: 2010, Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace, *Labour Economics* **17**(1), 276–283.

Besley, T. and Ghatak, M.: 2005, Competition and incentives with motivated agents, *American economic review* **95**(3), 616–636.

Besley, T. and Ghatak, M.: 2006, Sorting with motivated agents: implications for school competition and teacher incentives, *Journal of the European Economic Association* **4**(2-3), 404–414.

Bettinger, E. P.: 2012, Paying to learn: The effect of financial incentives on elementary school test scores, *Review of Economics and Statistics* **94**(3), 686–698.

Bhattacharya, S. and Guasch, J. L.: 1988, Heterogeneity, tournaments, and hierarchies, *The Journal of Political Economy* pp. 867–881.

Björk, B.-C. and Solomon, D.: 2013, The publishing delay in scholarly peer-reviewed journals, *Journal of Informetrics* **7**(4), 914–923.

Bloom, N. and Reenen, J. V.: 2011, Human resource management and productivity, *Handbook of labor economics* **4**, 1697–1767.

BMBF: 2002, Gesetz zur reform der professorenbesoldung. ProfBesReformG.
URL: *http://www.bmbf.de/pubRD/profbesreformg.pdf*

BMI: 2007, Bericht zum besoldungsrechtlichen vergaberahmen bei der professorenbesoldung nach § 35 abs. 5 bundesbesoldungsgesetz, *Technical report*, Bundesministerium des Innern. Stand 29 February 2008.
URL: *http://www.parlament-berlin.de/ados/16/WissForsch/vorgang/Bericht%20Professorenbesoldu*

Boly, A.: 2011, On the incentive effects of monitoring: evidence from the lab and the field, *Experimental Economics* **14**(2), 241–253.

Borjas, G. J. and Doran, K. B.: 2014, Which peers matter? the relative impacts of collaborators, colleagues, and competitors, *Technical report*, National Bureau of Economic Research.

Bresnahan, T. F. and Trajtenberg, M.: 1995, General purpose technologies 'engines of growth?', *Journal of Econometrics* **65**(1), 83–108.

Bull, C., Schotter, A. and Weigelt, K.: 1987, Tournaments and piece rates: An experimental study, *The Journal of Political Economy* pp. 1–33.

Bundesbeamtengesetz: 1985.
URL: *http://www.beamtengesetz.de/*

Carpenter, J., Matthews, P. H. and Schirm, J.: 2010, Tournaments and office politics: Evidence from a real effort experiment, *The American Economic Review* **100**(1), 504–517.

Chen, K.-P.: 2003, Sabotage in promotion tournaments, *Journal of Law, Economics, and Organization* **19**(1), 119–140.

Chiappori, P.-A., Galichon, A. and Salanié, B.: 2014, The roommate problem-is more stable than you think, *Technical report*, CESifo Working Paper.

Choo, E. and Siow, A.: 2006, Who marries whom and why, *Journal of Political Economy* **114**(1), 175–201.

Detmer, H. and Preissler, U.: 2004, Abenteuer w, strategien, risiken und chancen, *Forschung und Lehre* (6), 308–311.

Detmer, H. and Preissler, U.: 2005, Die neue professorenbesoldung, ein ueberblick, *Forschung und Lehre* (5), 256–258.

Detmer, H. and Preissler, U.: 2006, Die w-besoldung und ihre anwendung in den bundeslaendern, *Beitraege zur Hochschulforschung* **28**(2), 50–65.

DFG: 2014, Video portal on the excellence initiative.
URL: *http://www.excellence-initiative.com/excellence-initiative*

Diamantoudi, E., Miyagawa, E. and Xue, L.: 2004, Random paths to stability in the roommate problem, *Games and Economic Behavior* **48**(1), 18–28.

Dickinson, D. L.: 1999, An experimental examination of labor supply and work intensities, *Journal of Labor Economics* **17**(4), 638–670.

Dickinson, D. and Villeval, M.-C.: 2008, Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories, *Games and Economic behavior* **63**(1), 56–76.

Dilger, A.: 2013, Vor- und nachteile der w-besoldung, *Discussion Paper of the Institute for Organisational Economics* . Westfaelische Wilhelms-Universitaet Muenster.
URL: *http://www.econstor.eu/bitstream/10419/74542/1/747328412.pdf*

Dohmen, T. and Falk, A.: 2011, Performance pay and multidimensional sorting: Productivity, preferences, and gender, *The American Economic Review* pp. 556–590.

Expertenkommission: 2000, Bericht der expertenkommission - reform des hochschuldienstrechts, *Technical report*.
URL: *http://www.bmbf.de/pub/Bericht_.pdf*

Filmuniversitaet: 2014.
URL: *http://www.filmuniversitaet.de/de/filmuniversitaet.html*

Fitzenberger, B. and Schulze, U.: 2014, Up or out: Research incentives and career prospects of postdocs in germany, *German Economic Review* **15**(2), 287–328.

Forschung und Lehre: 1999-2013.

URL: *http://www.forschung-und-lehre.de/wordpress/?page_id=7*

Forschung und Lehre, "Wir ueber uns": 2014.

URL: *http://www.forschung-und-lehre.de/wordpress/?page_id=29*

Freeman, R. B. and Gelber, A. M.: 2010, Prize structure and information in tournaments: Experimental evidence, *American Economic Journal: Applied Economics* **2**(1), 149–164.

Frey, B. S. and Jegen, R.: 2001, Motivation crowding theory, *Journal of Economic Surveys* **15**(5), 589–611.

FuL: 2002, Habilitationen und berufungen, *Forschung und Lehre* p. 41.

Gale, D. and Shapley, L. S.: 1962, College admissions and the stability of marriage, *American Mathematical Monthly* pp. 9–15.

Georgellis, Y., Iossa, E. and Tabvuma, V.: 2010, Crowding out intrinsic motivation in the public sector, *Journal of Public Administration Research and Theory* p. muq073.

Glewwe, P., Ilias, N. and Kremer, M.: 2010, Teacher incentives, *American Economic Journal: Applied Economics* pp. 205–227.

Grigoriou, K. and Rothaermel, F. T.: 2014, Structural microfoundations of innovation the role of relational stars, *Journal of Management* **40**(2), 586–615.

Griliches, Z.: 1958, Research costs and social returns: Hybrid corn and related innovations, *The Journal of Political Economy* pp. 419–431.

Griliches, Z.: 1998, The search for r&d spillovers, *R&D and Productivity: The Econometric Evidence*, University of Chicago Press, pp. 251–268.

Gürtler, O. and Kräkel, M.: 2010, Optimal tournament contracts for heterogeneous workers, *Journal of Economic Behavior & Organization* **75**(2), 180–191.

Haeck, C. and Verboven, F.: 2012, The internal economics of a university: evidence from personnel data, *Journal of Labor Economics* **30**(3), 591–626.

Hajduková, J.: 2006, Coalition formation games: A survey, *International Game Theory Review* **8**(04), 613–641.

Handel, K. C.: 2005, *Die Umsetzung der Professorenbesoldungsreform in den Bundesländern*, CHE.

URL: *http://www.che.de/downloads/Profbesoldung_Laendervergleich_AP65.pdf*

Harbring, C. and Irlenbusch, B.: 2008, How many winners are good to have?: On tournaments with sabotage, *Journal of Economic Behavior & Organization* **65**(3), 682–702.

Harbring, C. and Irlenbusch, B.: 2011, Sabotage in tournaments: Evidence from a laboratory experiment, *Management Science* **57**(4), 611–627.

Harbring, C., Irlenbusch, B. and Kräkel, M.: 2004, Ökonomische analyse der professorenbesoldungsreform in deutschland, *Bildung–33. Wirtschaftswissenschaftliches Seminar Ottobeuren. Tübingen: Mohr Siebeck*, pp. 197–219.

Historie 1872 bis heute: 2014.

URL: *http://www.hs-geisenheim.de/hochschule/geschichtezahlen.html*

Hochschulkompass: 2014.

URL: *http://www.hochschulkompass.de/hochschulen/hochschulen-in-deutschland-die-hochschulsuche.html*

Hochschullehrerbund: 2009, Die professorengehaelter in der w-besoldung, art und umfang von berufungsverhandlungern, Online.

URL: *http://professur.fh-hannover.de/bewerbung/Hochschullehrerbund_Leistungsbezuege.pdf*

Hossain, T. and List, J. A.: 2012, The behavioralist visits the factory: Increasing productivity using simple framing manipulations, *Management Science* **58**(12), 2151–2167.

Jaffe, A. B., Trajtenberg, M. and Henderson, R.: 1993, Geographic localization of knowledge spillovers as evidenced by patent citations, *The Quarterly Journal of Economics* **108**(3), 577–598.

Jones, B. F.: 2009, The burden of knowledge and the death of the renaissance man: is innovation getting harder?, *The Review of Economic Studies* **76**(1), 283–317.

Jongbloed, B.: 2009, *Higher Education Funding Systems: An Overview covering 5 European jurisdictions and the Canadian province of Ontario*, Center for Higher Education Policy Studies (CHEPS), chapter 2, pp. 8–15.

Journal Citation Report: 2000-2012.

Kaiser, F., Vossensteijn, H. and Koelman, J.: 2002, *Public funding of higher education - A comparative study of funding; Mechanisms in 10 countries*, Center for Higher Education Policy Studies (CHEPS), chapter 7, pp. 69–76.
URL: *doc.utwente.nl/45235/1/engreport02publicfunding.pdf*

Kantor, S. and Whalley, A.: 2014, Knowledge spillovers from research universities: evidence from endowment value shocks, *Review of Economics and Statistics* **96**(1), 171–188.

Kim, E. H., Morse, A. and Zingales, L.: 2009, Are elite universities losing their competitive edge?, *Journal of Financial Economics* **93**(3), 353–381.

Klaus, B., Klijn, F. and Walzl, M.: 2010, Stochastic stability for roommate markets, *Journal of Economic Theory* **145**(6), 2218–2240.

Kräkel, M.: 2003, U-type versus j-type tournaments as alternative solutions to the unverifiability problem, *Labour Economics* **10**(3), 359–380.

Kräkel, M.: 2006, Zur reform der professorenbesoldung in deutschland, *Perspektiven der Wirtschaftspolitik* **7**(1), 105–126.

Kräkel, M. and Sliwka, D.: 2004, Risk taking in asymmetric tournaments, *German Economic Review* **5**(1), 103–116.

Kremer, M.: 1993, The o-ring theory of economic development, *The Quarterly Journal of Economics* **108**(3), 551–575.

Kuerschners Deutscher Gelehrten Kalender: 2006, cd-rom. Data set used for 21st edition but containing more and more comprehensive data.

Kuerschners Deutscher Gelehrten Kalender: 2008, cd-rom. Data set used for 22nd edition but containing more and more comprehensive data.

Kuerschners Deutscher Gelehrten Kalender Online: 2013.
URL: *http://www.degruyter.com/databasecontent?dbid=kdgo&dbsource=%2Fdb%2Fkdgo*

Lach, S. and Schankerman, M.: 2004, Royalty sharing and technology licensing in universities, *Journal of the European Economic Association* **2**(2-3), 252–264.

Lach, S. and Schankerman, M.: 2008, Incentives and invention in universities, *The RAND Journal of Economics* **39**(2), 403–433.

Lavy, V.: 2009, Performance pay and teachers' effort, productivity, and grading ethics, *The American Economic Review* pp. 1979–2011.

Lazear, E. P.: 2000a, Performance pay and productivity, *The American Economic Review* **90**(5), 1346–1361.

Lazear, E. P.: 2000b, The power of incentives, *American Economic Review* pp. 410–414.

Lazear, E. P.: 2004, The peter principle: A theory of decline, *Journal of Political Economy* **112**(S1), S141–S163.

Lazear, E. P. and Oyer, P.: 2012, Personnel economics, *The Handbook of Organizational Economics* p. 479.

Lazear, E. P. and Rosen, S.: 1981, Rank-order tournaments as optimum labor contracts, *The Journal of Political Economy* **89**(5), 841–864.

Lazear, E. P. and Shaw, K. L.: 2007, Personnel economics: The economist's view of human resources, *The Journal of Economic Perspectives* pp. 91–114.

Legros, P. and Newman, A. F.: 2002, Monotone matching in perfect and imperfect worlds, *The Review of Economic Studies* **69**(4), 925–942.

Lemieux, T., Macleod, W. B. and Parent, D.: 2009, Performance pay and wage inequality, *The Quarterly Journal of Economics* **74**(1), 1–49.

Leuven, E., Oosterbeek, H., Sonnemans, J. and Van Der Klaauw, B.: 2011, Incentives versus sorting in tournaments: Evidence from a field experiment, *Journal of Labor Economics* **29**(3), 637–658.

Lucas, R. E.: 1988, On the mechanisms of economic development, *Journal of Monetary Economics* **22**(1), 3–32.

Lucas, R. E.: 1990, Why doesn't capital flow from rich to poor countries?, *The American Economic Review* **80**(2), 92–96.

Lünstroth, P.: 2011, *Leistungslohn und Kooptation-Eine ökonomische Analyse der Reform der Professorenbesoldung*, PhD thesis, Universitaet Trier.
URL: *http://ubt.opus.hbz-nrw.de/volltexte/2012/739/pdf/Dissertation_Pia_Luenstroth.pdf*

McCormack, J., Propper, C. and Smith, S.: 2014, Herding cats? management and university performance, *The Economic Journal* **124**(578), F534–F564.

McFadden, D.: 1974, Conditional logit analysis of qualitative choice behavior, *Frontiers in Econometrics* pp. 105–142.

Mohr, J.: 2007, Professoren, die vertreibung der weisen, *Spiegel Online* .
URL: *http://www.spiegel.de/unispiegel/jobundberuf/professoren-die-vertreibung-der-weisen-a-458419-druck.html*

Muralidharan, K. and Sundararaman, V.: 2011, Teacher performance pay: Experimental evidence from india, *The Journal of Political Economy* **119**(1), 39–77.

NationalScienceBoard: 2000, Science and engineering indicators. National Science Foundation, Washington D.C.

Oeffentlicher-Dienst: 2004, Gesetz ueber die erhoehung von dienst- und versorgungsbezuegen in bund und laendern 2003/2004. The parent page of the website, http://oeffentlicher-dienst.info/beamte/bund/a/, contains salary tables from 1927 till now (latest 01.03.2015-29.02.2016). Last checked 11-08.2015.
URL: *http://oeffentlicher-dienst.info/c/t/rechner/beamte/bund/a/2004i?id=beamte-bund-2004i&matrix=1*

Oettl, A.: 2012, Reconceptualizing stars: Scientist helpfulness and peer performance, *Management Science* **58**(6), 1122–1140.

Oyer, P. and Schaefer, S.: 2011, Personnel economics: Hiring and incentives, *Handbook of Labor Economics* **4**, 1769–1823.

"Personalmittelbudgetierung, Empfehlungen zu ihrer Ausgestaltung": 2008.
URL: *http://www.uni-kanzler.de/fileadmin/Dateien/Pressemitteilung_Personalmittelbudgetierung.p*

Phelps, C., Heidl, R. and Wadhwa, A.: 2012, Knowledge, networks, and knowledge networks a review and research agenda, *Journal of Management* **38**(4), 1115–1166.

Preissler, U.: 2006, Zwischenbilanz professorenbesoldung - sichtweise und beratung des deutschen hochschulverbandes. Arbeitsgruppe Fortbildung im Sprecherkreis der Deutschen Universitätskanzler.
URL: *https://www.uni-muenster.de/.../prei_ler_jena_5.06.doc*

Pritchard, R.: 2006, Trends in the restructuring of german universities, *Comparative Education Review* **50**(1), 90–112.

Research and Academic Jobs in Germany: 2011.
URL: *http://www.careeredu.eu/index.php?doc=1*

Romer, P. M.: 1986, Increasing returns and long-run growth, *The journal of political economy* **94**(5), 1002–1037.

Romer, P. M.: 1990, Endogenous technological change, *Journal of political Economy* **95**(8), S71–S102.

Roth, A. E. and Vate, J. H. V.: 1990, Random paths to stability in two-sided matching, *Econometrica: Journal of the Econometric Society* pp. 1475–1480.

Schniederjuergen, A.: 2013a, E-mail.

Schniederjuergen, A.: 2013b, E-mail.

Schotter, A. and Weigelt, K.: 1992, Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results, *The Quarterly Journal of Economics* pp. 511–539.

Shapley, L. S. and Shubik, M.: 1971, The assignment game i: The core, *International Journal of Game Theory* **1**(1), 111–130.

Shearer, B.: 2004, Piece rates, fixed wages and incentives: Evidence from a field experiment, *The Review of Economic Studies* **71**(2), 513–534.

Shi, L.: 2010, Incentive effect of piece-rate contracts: Evidence from two small field experiments, *The BE Journal of Economic Analysis & Policy* **10**(1).

Singh, J.: 2005, Collaborative networks as determinants of knowledge diffusion patterns, *Management science* **51**(5), 756–770.

Siow, A.: 2009, Testing becker's theory of positive assortative matching, *University of Toronto WP* **356**.

Vossensteyn, J. and Jongbloed, B.W.A. abd Koelman, J.: 1998, *University Funding Mechanisms and related issues*, Ministerio dell 'universita e della ricerca scientifica e tecnologia, chapter 5, pp. 41–53.

Waldinger, F.: 2010, Quality matters: The expulsion of professors and the consequences for phd student outcomes in nazi germany, *Journal of Political Economy* **118**(4), 787–831.

Waldinger, F.: 2012, Peer effects in science: Evidence from the dismissal of scientists in nazi germany, *The Review of Economic Studies* **79**(2), 838–861.

Wuchty, S., Jones, B. F. and Uzzi, B.: 2007, The increasing dominance of teams in production of knowledge, *Science* **316**(5827), 1036–1039.

Ytsma, E.: 2015, Lone stars or constellations? the impact of performance pay on matching assortativeness in academia.