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

Erina Ytsma†, LSE

30 November 2014

Abstract

This paper studies the effect of performance related pay on matching assortativeness. Positive assortativeness by worker productivity increases total output if there are complementarities in worker skill (Legros and Newman 2002), while such complementarities in turn cause matching to be positive assortative[1] (Kremer 1993). I present a simple matching model in which there are positive spillovers from worker productivity, but worker utility also depends on idiosyncratic preferences regarding colleagues and location of work. The positive spillovers cause matching to be positively assortative, but the frictions cause assortativeness to decrease. Introducing performance related pay increases positive assortative matching again and this increase is larger when complementarities are stronger. I test this hypothesis empirically 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 to analyse changes in faculty composition over time. I find that matching becomes more assortative after the reform and the increase is economically large. I test whether this increase is larger if complementarities are stronger in a diff-in-diff framework using the average number of authors on a paper as a proxy for the strength of complementarities in a field and find that assortative matching increases more in fields with stronger complementarities. The paper thus provides empirical evidence that performance related pay increases positive assortative matching. This result is robust to controlling for alternative explanations, most notably pre-existing trends and differential hiring budget.

JEL classification: J33, M51, M52

*I would like to thank Oriana Bandiera, Tim Besley and Andrea Prat for their continued advice and support. I would also like to thank Steve Pischke, Mark Schankerman, Axel Schniederjuergen, Baran Duzce, De Gruyter, the ministries of education of the German states and seminar participants at LSE and STICERD for helpful comments, information or data.

†e.g.ytsma@lse.ac.uk

[1] Specifically, if output (and payoff) functions are supermodular in skill.

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 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[2], 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 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 paper 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 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 positive assortative matching by productivity again, and this increase is larger if complementarities are stronger. I test this hypothesis empirically using the introduction of performance pay in German academia as a natural experiment and analysing its effect on the matching assortativeness of academics.

A reform introducing a new professorial salary scheme comprising a basic wage plus performance related bonuses was passed by the German parliament in 2002. The new performance related pay scheme replaced the old pay scheme in which professorial salaries increased with age (Handel 2005). German states had until 1 January 2005 to implement the reform within their respective jurisdiction (Detmer and Preissler 2005). 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 nation-wide 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. Academics that 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 under the new performance pay scheme and therefore under the influence of incentives to match more or less assortatively that come with performance pay.

For the empirical analysis I constructed a new data set comprising the affiliations, productivity and related information of the universe of German academics. From this data set I derived a department level data set that I use to study any changes in departmental composition from before to after the reform. The combination

[2] See Lazear (1999), Bloom and Reenen (2011), Lazear and Shaw (2007) for an overview of the literature.

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].

I first assess whether positive assortative matching increases post-reform and subsequently test whether any increase is caused by the introduction of performance pay. 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 if performance pay increases positive assortative matching by comparing the difference in changes in positive assortativeness in fields with weaker and stronger complementarities in a difference in differences framework, 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 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.

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 that 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 that retire in a year and this variation is historically determined and plausibly exogenous.

To my knowledge this paper is one of the first to study the effect of performance pay on the level of assortative matching. Bandiera et al. (2013) study the effect of team-based incentives on matching assortativeness in teams within a firm employing manual labourers. This paper 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. This setting is interesting and relevant, because 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 PhDs (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 paper 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 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

[3] To my knowledge, this is the first paper to study the German academic pay reform using a data set that encompasses the universe of German academics.

[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.

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, 1986) and Lucas (1990, 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 matter for total welfare. A 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 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 in which the distribution of the production factors (and output) is an important determinant of welfare, such as academia.

The rest of the paper is organised as follows; I provide the institutional background in section 2, present the theoretical framework in section 3, describe the data in section 4 and discuss the empirical results in section 6. Section 7 concludes.

2 Institutional Background

2.1 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 paper (*Hochschulkompass* 2014)[6]. 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 (Frans Kaiser 2002). The public subsidies to universities have been traditionally

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

[6] 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)).

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.

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 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, that then decides which of the candidates on the list to appoint. (Lünstroth 2011).

2.3 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), 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.

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 exceptional teaching evaluations can for instance 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 euro per month (e.g. Bielefeld and Bremen University) (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 somewhere between 200 and 600 euro (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)[7]. 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 euro per month upon positive evaluation (Detmer and Preissler 2005)[8]. This is only a very small bonus compared to the total amount of bonuses that tenured professors may get awarded, which can go up to 5241,48 euro[9], or more in special circumstances[10](BMBF 2002, Detmer and Preissler 2005). Importantly, universities have discretion on how to award bonuses [11].

The academic pay reform includes a cost-neutrality clause that stipulates that the average professorial

[7] 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)

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

[9] This limit 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)

[10] 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).

[11] 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.

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-to-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 euro at universities, about 26% of this is available for performance pay bonuses for university professors[12]. 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 the respective university (Handel 2005). The fact that the 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 bench mark professorial pay average is reasonably binding each year (Harbring et al. 2004, Kräkel 2006, Lünstroth 2011).

3 Model

In this section I present a simple one-sided matching model of team (in the context of the paper: 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 that 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 static 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 idiosyncratic 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.[13] 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 only if they are friends, whereas it depends positively on partner's productivity regardless of their relationship in this model due to positive spillovers. The model in Bandiera et al. (2013) also comprises a component that does not depend on agents' productivity, but in contrast to the idiosyncratic component in this paper, 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

[12] 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.

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

of agents so as to allow for any kind of unobservable preferences regarding potential matching partners' identity[14].

3.1 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:

$$u_{i|j} = \begin{cases} \alpha_i f\left(\theta_i \mid \theta_j\right) + w_{i|j} & \text{if } i \text{ is matched with academic } j \\ \underline{u}_i & \text{if } i \text{ is unmatched} \end{cases} \tag{1}$$

where $f\left(\theta_i \mid \theta_j\right)$ is the productivity of i when he is matched with j, α_i is an individual specific weighing constant that captures i's valuation of this productivity (for instance capturing his 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, $\underline{u}_i = \underline{u}, \forall i$.

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 [15]. The model is intentionally kept sufficiently general to allow for spillovers that are not restricted to co-authorships, but simply occur between academics that are working in the same department.

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

$$\mu\left(i\right) = j \text{ if } i \text{ is matched with academic } j \tag{2}$$

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 $u_{i|\mu(i)} = \alpha_i f\left(\theta_i \mid \theta_{\mu(i)}\right) + w_i \geq \underline{u}, \forall i \in \{1, .., m\}$. A matching μ cannot be improved upon if there are no two academics i, j such that

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

and

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

[14] This is because affiliation decisions are likely influenced by more than just friendship ties, such as the geographical location of a department.

[15] 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 $\theta_j > \theta_k$

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:$\sum_i u_{i|\mu(i)} \geq \sum_i u_{i|\mu'(i)}$.

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*[16]. (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 that 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 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[17]. 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[18]. I will show in section 3.3 that performance pay is such an innovation.

3.2 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, etc.), but this benefit is not related to an academic's

[16] 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 $f\left(\theta_i' \mid \theta_j'\right) - f\left(\theta_i \mid \theta_j'\right) \geq f\left(\theta_i' \mid \theta_j\right) - f\left(\theta_i \mid \theta_j\right)$

[17] In expectation, if no stable matching exists

[18] If the academic output function exhibits increasing differences

salary. The α_i 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 $\alpha_i f(\theta_i \mid \theta_j)$, 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. I will make the following assumptions regarding the distribution of this noise term:

- *assumption*1.a[19] : ν_{ij} are random and independent draws from $\nu \sim N\left(0, \sigma^2\right)$
- *assumption*1.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*[20]*. This probability is smaller if complementarities are stronger*[21]*.* (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'.

If the academic production function $f(\theta_i \mid \theta_j)$ exhibits increasing differences, proposition 2 implies that total academic output is not maximised with non-zero probability if academics have idiosyncratic personal preferences regarding potential matching partners. 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 $f(\theta_i \mid \theta_j)$ exhibits increasing differences, reduce the probability with which total academic output is less than maximal.

3.3 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.

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 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. Whether the former exactly cancels the latter depends on the functional form of the academic productivity

[19] 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 paper however only derives from the change in pay scheme, which in turn changes the utility from own academic productivity. Therefore, I can let the noise distribution be more general.

[20] *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).*

[21] Complementarities are stronger if $\partial f(\theta_i \mid \theta_j) / \partial \theta_j$ is larger

function $f(\theta_i \mid \theta_j)$[22]. I refrain from making assumptions regarding the specification of the academic output function here, because I do not know of stylised facts that could inform any such assumptions.

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_i 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 $\alpha_i^{post} > \alpha_i^{pre}$), 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 between fields with differing 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.

4 Data Description

For this project I built a new panel data set comprising 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

[22] As well as the relative sizes of the α_i's and tournament bonuses.

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 that 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.

For the purposes of this research, I focus on academics who held a tenured position at a German public university between 1999 and 2013. I focus on these academics, 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[23]. 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.

I derived the department panel data set from a panel data set of the universe of German academics for the years 1993-2013 that I constructed for this project. This individual level panel comprises 55132 academics that had a tenured position at a German public university between 1999 and 2013[24]. 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 previous 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.

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

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

[24] 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.

sciences, nutrition and veterinary medicine; medicine (human); dentistry[25].

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[26].

I can match 83% of academics that 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[27] are 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[28]. 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 Appendix B.

All in all, the data set used for this paper 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[29].

4.1 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[30].The average productivity of new hires has a mean[31] of 3.68 and a standard deviation of 8.68[32](cf. Table 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 that 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 that 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 depart-

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

[26] 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.

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

[28] 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.

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

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

[31] 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

[32] 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.

ment as a measure of the quality of a department[33]. The mean of the department quality variable is 2.22 and the standard deviation is 4.15. I use the number of tenured academics that 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 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 Empirical Analysis

The main aim of the paper 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 or 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[34] 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.

5.1 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 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 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. Even so, 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[35].

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

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

[35] 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

As a first check for changes in assortatitiveness, 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. Figures 2A, 2B and 2C 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 figures 3A through 3C and 4A through 4C 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.

Figures 5A, 5B and 5C show the pre- and post-reform[36] kernel density plots of 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.

5.2 Regression analysis

To formally test whether positive assortative matching increases post-reform, I estimate the following random effects panel data model[37] for department j in field f and year t:

$$\bar{y}_{j,f,t}^{\{k\}} = \beta_0 + \beta_1 \bar{y}_j^{old} + \beta_2 t + \beta_3 post + \beta_4 post \cdot \bar{y}_j^{old} + c_f + u_{jt} \tag{3}$$

The dependent variable $\bar{y}_{j,t}^{\{k\}}$ 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 $\bar{y}_j^{old}$; the average productivity of the affiliates of department j in 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 $post \cdot \bar{y}_j^{old}$ is the interaction of this variable with department quality. The c_f are field fixed effects[38].

from outside Germany in order to analyse their average productivity and sorting patterns in to German departments in the future.

[36] The sample is not restricted to the year before and after the reform only for these plots, since effort changes due to the reform would actually go against the pattern expected for leavers if there is an increase in positive assortative matching.

[37] 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.

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

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, 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[39] and to minimise any bias introduced by reforms and other events taking place around the same time as the reform[40]. 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 2. Throughout the paper 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 trend, while a linear time trend is added in column b.

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

[39] 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.

[40] 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.

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 (2010, 2012) that spillover effect are largest for early-career academics [41].

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 significantly 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.

5.2.1 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 :

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

$$+\beta_6 post \cdot \bar{y}_j^{old} + \beta_7 post \cdot Complementarity_f + \beta_8 post \cdot Complementarity_f \cdot \bar{y}_j^{old} + c_f + u_{jt} \qquad (4)$$

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

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

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

The estimation results of regression (4) for new hires, junior hires, senior hires and leavers form the core results of this paper and can be found in Table 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 (2010, 2012) that spillover effect are largest for early-career academics. The positive and significant interaction of departmental quality and field complementarity strength in the leavers regressions (columns 4a and 4b) also aligns with matching being more positive assortative in high complementarity fields, as it indicates that the difference in average productivity of former affiliates between high and low quality departments is larger in high complementarity fields.

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.

5.2.2 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.

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:

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

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

$$+\beta_{11} post'04 \cdot Complementarity_f \cdot \bar{y}_j^{old} + \beta_{12} post \cdot Complementarity_f \cdot \bar{y}_j^{old} + c_f + u_{jt} \quad (5)$$

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 4). If there are preexisting 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[43], 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[44] with department quality and field complementarity strength smaller.

Table 4 shows the estimation results of the specification in (5). 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 3), especially for junior hires. Moreover, the placebo-post dummy triple interactions are not significant in any of these regressions, and the coefficients are 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 :

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

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

$$+\beta_{11} post'0\{l\} \cdot Complementarity_f \cdot \bar{y}_j^{old} + \beta_{12} post \cdot Complementarity_f \cdot \bar{y}_j^{old} + c_f + u_{jt}, l \in [2,3,4] \quad (6)$$

[43] 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.

[44] Labelled $post'05$ in Table 4 for extra clarity.

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. The results of these regressions are very similar to the pre-trend regressions reported in Table 4; the triple interactions of $post'05$ with department quality and field complementarity strength remains positive and significant for all new hires and junior hires and the coefficients have similar sizes as in the baseline regression (Table 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.

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 are thus able to 'buy' 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[45] Post-reform however, universities can offer bonuses to attract academics or prevent affiliates from moving to another university[46]. 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 that 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 (Frans Kaiser 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).

[45] 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).

[46] 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

Because a professor that 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 that retire from a department in a given year is historically determined and should not be correlated with contemporaneous factors[47].

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 (J.J. Vossensteyn 1998, Frans Kaiser 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[48] as a robustness check later on.

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

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

$$+\beta_7 post \cdot \bar{y}^{old}_j + \beta_8 post \cdot Complementarity_f + \beta_9 post \cdot Complementarity_f \cdot \bar{y}^{old}_j + \beta_{10} post \cdot B_{j,t} + c_f + u_{jt} \quad (7)$$

The variable $B_{j,(t,t-1)}$ is the number of professors that retire (turn 66) between t and $t-1$ 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[49].

The estimation results of this specification are reported in Table 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 3 with the results in Table 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

[47] 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 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 that 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)

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

[49] 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 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.

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 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 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), pp. 55-56)[50]. An academic that 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 that 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 retainment 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 retainment 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

[50] 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.

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[51].

Robustness Checks

Finally, I do a number of robustness checks, the results of which are reported in Table 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 4) as a fixed effects panel data model[52]. The estimation results barely change compared to the baseline regression in Table 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 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 than the baseline results in Table 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 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, (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 7) 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 disposable to higher rank departments in high complementarity fields, this could

[51] Note that supermodularity of the academic output function could also explain why the interaction of the *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.

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

cause the triple interaction in the baseline regression to be positive. The regressions in table 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 that 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 3 and 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.

6 Conclusion

This paper 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 be able to study the effect of the performance pay on matching assortativeness, I use the introduction of a performance pay scheme in German academia as a natural experiment and I constructed a new data set of the affiliations and productivity of the universe of German academics. 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 sectorwide).

The paper 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 post-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 paper 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.

References

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

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.

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.: 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.

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*

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.

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

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.

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*

Frans Kaiser, Hans Vossensteijn, J. K.: 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*

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.

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

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.

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.

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.

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

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.

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.

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.

Lazear, E. P.: 1999, Personnel economics: past lessons and future directions, *Technical report*, National bureau of economic research.

Lazear, E. P. and Shaw, K. L.: 2007, Personnel economics: The economist's view of human resources, *Technical report*, National Bureau of Economic Research.

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.

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*

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*

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

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

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.

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.

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

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

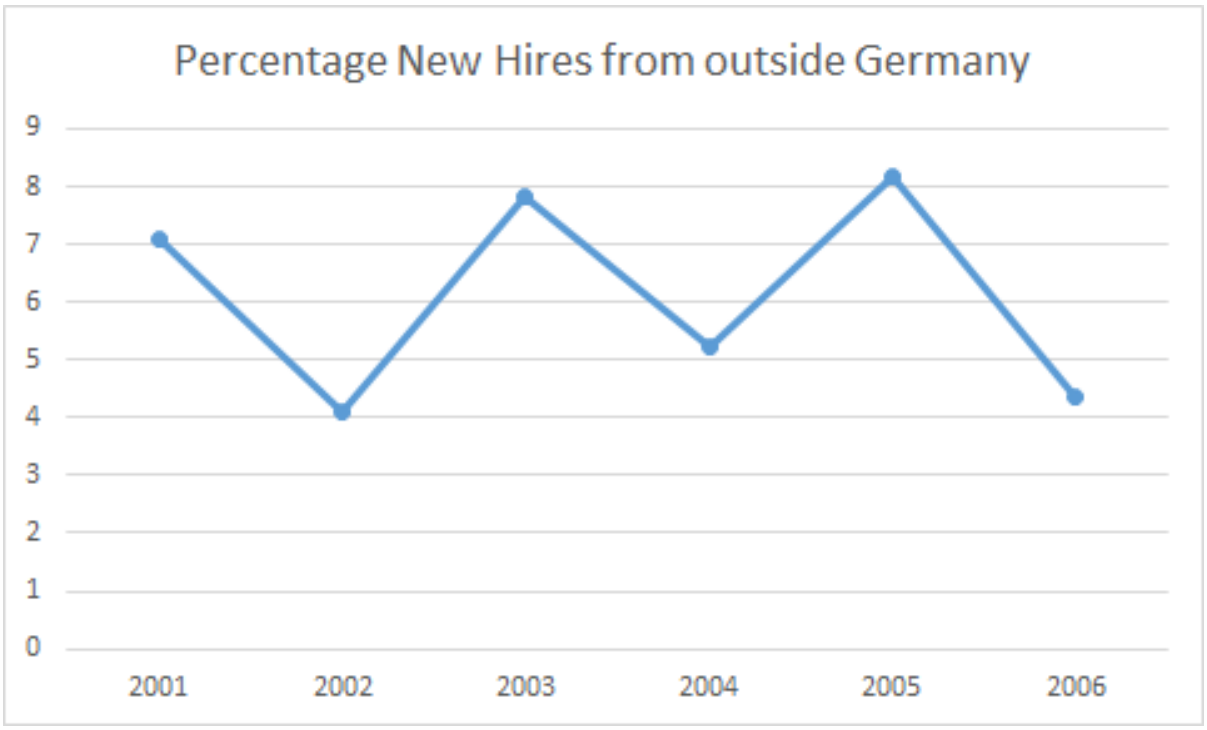
Siow, A.: 2009, Testing becker's theory of positive assortative matching, *University of Toronto WP* **356**.

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.

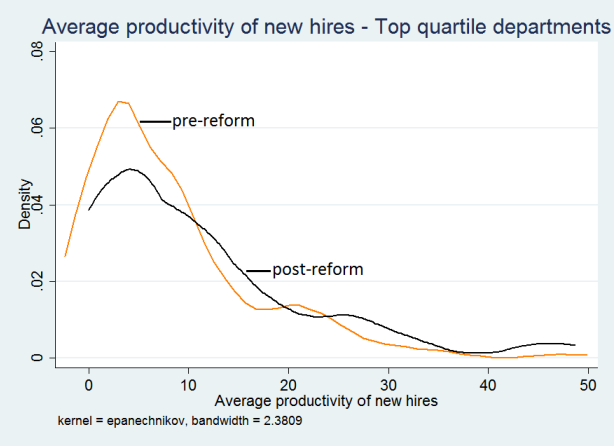
Fig. 1



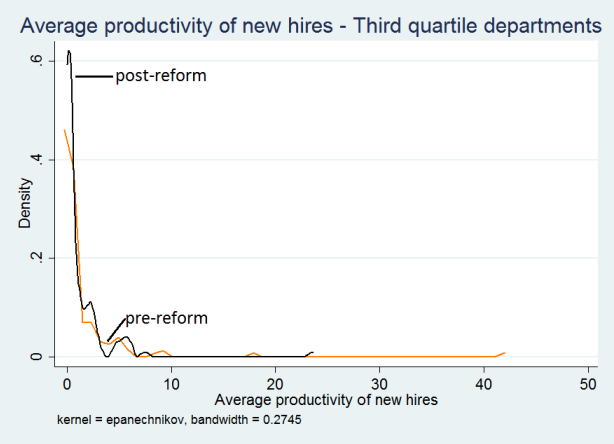
The figure above shows the number of new hires attracted from outside Germany as a percentage of the total number of new hires in a given year for the years 2001-2006.

Fig. 2

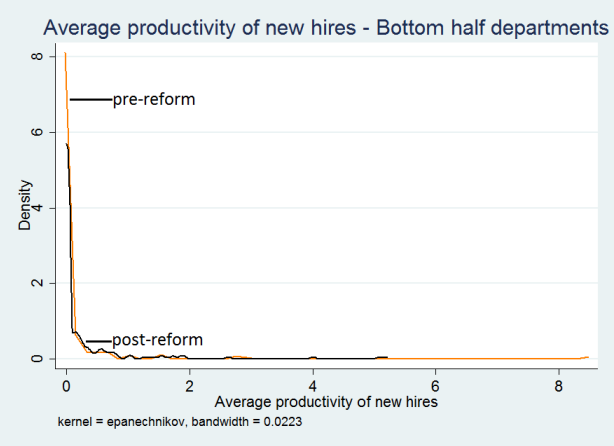
(a)



(b)



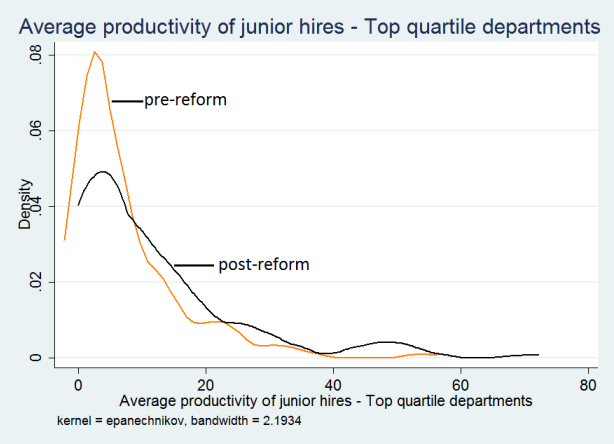
(c)



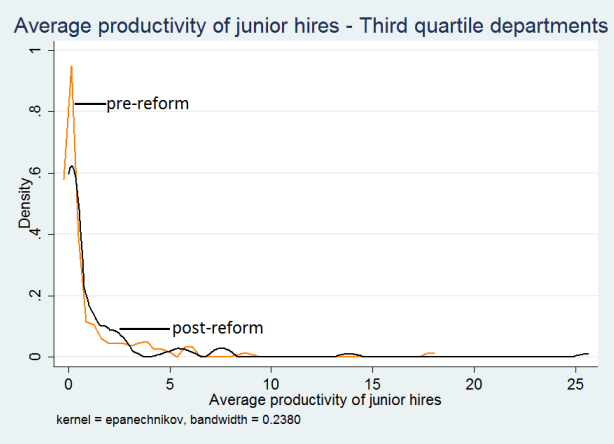
The above figures are kernel density plots of the average productivity of new hires in the year before the reform came into effect (2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department 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.

Fig. 3

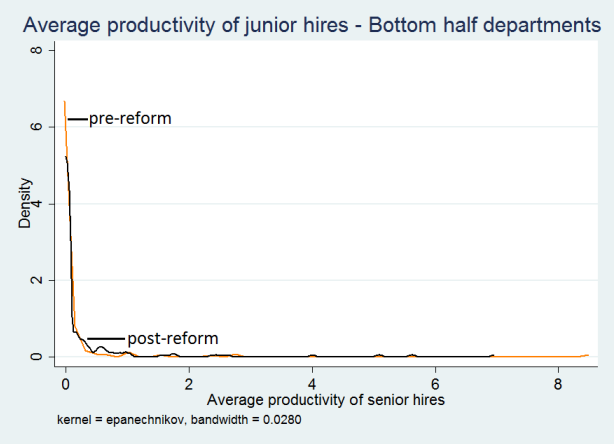
(a)



(b)



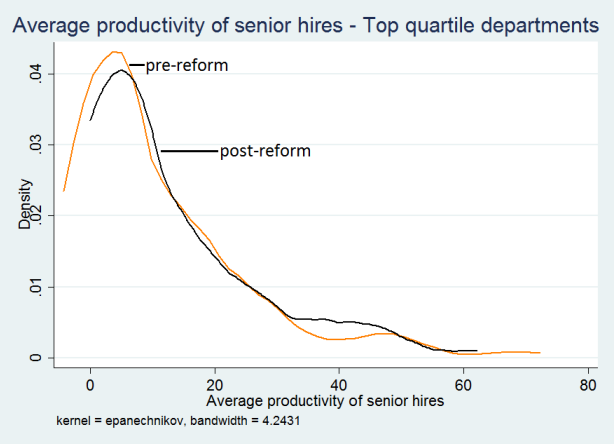
(c)



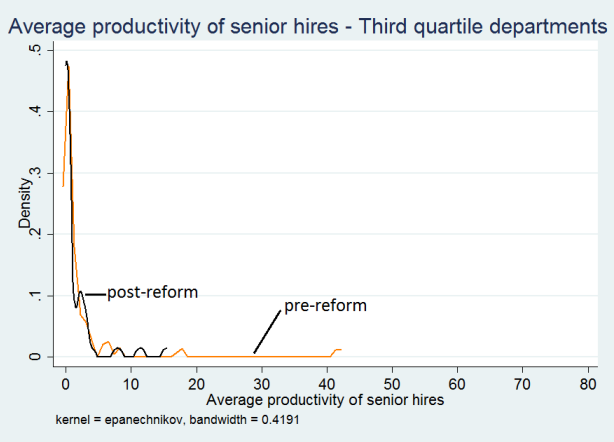
The above figures are kernel density plots of the average productivity of junior hires in the year before the reform came into effect 2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department 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.

Fig. 4

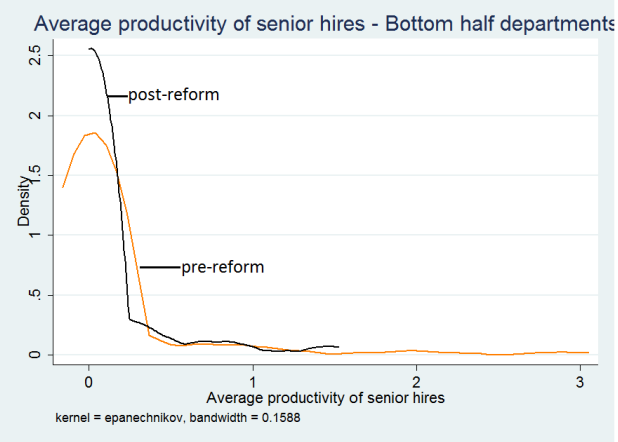
(a)



(b)



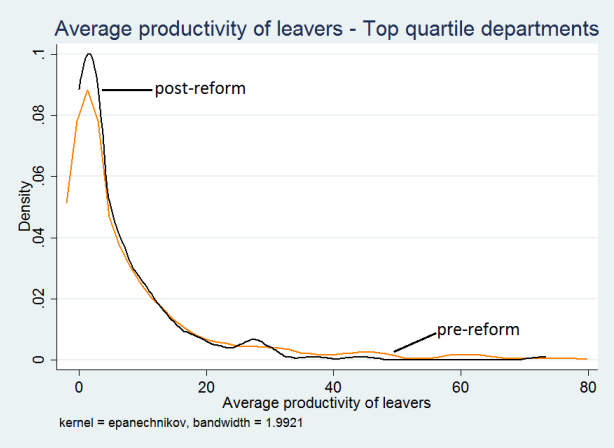
(c)



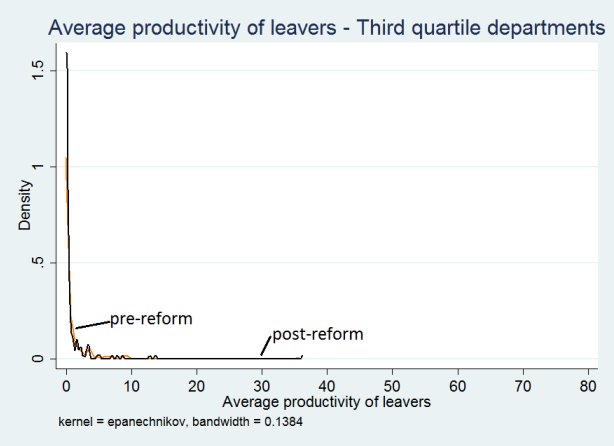
The above figures are kernel density plots of the average productivity of junior hires in the year before the reform came into effect 2004) and the year after (2006), for the departments in, respectively, the top quartile, the second quartile from the top and the lowest two quartiles of the department 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.

Fig. 5

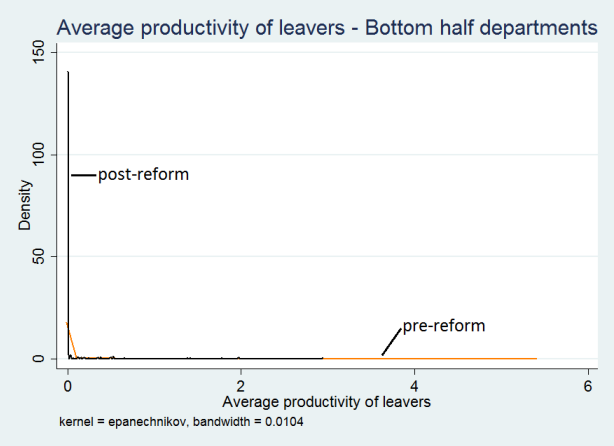
(a)



(b)



(c)



The above figures are kernel density plots of the average productivity of academics that 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.

Tab. 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

Tab. 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.

Tab. 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.

Tab. 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.

Tab. 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.

Tab. 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.

Tab. 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:

$$\tilde{\mu} = \begin{cases} \tilde{\mu}(1) = 2, \tilde{\mu}(3) = 4, .., \tilde{\mu}(k+2) = k+1, \tilde{\mu}(k) = k, .., \tilde{\mu}(m) = m & \text{if } k \text{ odd} \\ \tilde{\mu}(1) = 2, \tilde{\mu}(3) = 4, .., \tilde{\mu}(k+1) = k+1, \tilde{\mu}(k) = k, .., \tilde{\mu}(m) = m & \text{if } k \text{ even} \end{cases} \tag{8}$$

where:

$$k = \begin{cases} i : u_{k|k+1} < \underline{u}, u_{k+1|k} \geq \underline{u} & \text{for } i \text{ even} \\ i : u_{k|k-1} < \underline{u}, u_{k+1|k+2} \geq \underline{u} & \text{for } i \text{ odd} \end{cases} \tag{9}$$

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*.

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 $D^{\nu}_{k(i,j)} = \nu_{k,i} - \nu_{k,j}$. Then i and $i+t$, with $t \notin \{0,1\}$ form a blocking pair if $\alpha_i f(\theta_i \mid \theta_{i+t}) - \nu_{i,i+t} \geq \alpha_i f(\theta_i \mid \theta_{i+1}) - \nu_{i,i+1}$ and $\alpha_{i+t} f(\theta_{i+t} \mid \theta_i) - \nu_{i+t,i} \geq \alpha_{i+t} f\left(\theta_{i+t} \mid \theta_{\tilde{\mu}(i+t)}\right) - \nu_{i+t,\tilde{\mu}(i+t)}$. That is, $\tilde{\mu}$ is not stable if:

$$\alpha_i D^{\theta}_{i(i+1,i+t)} \leq D^{\nu}_{i(i+1,i+t)} \tag{10}$$

and

$$\alpha_{i+t} D^{\theta}_{i+t(i,\tilde{\mu}(i+t))} \geq D^{\nu}_{i+t(i,\tilde{\mu}(i+t))} \tag{11}$$

with $t \notin \{0,1\}$ and at least one inequality strict. By assumptions 1.*a* and 1.*b*, this happens with non-zero probability.

The probability that the matching $\tilde{\mu}$ is not stable is smaller if complementarities are larger. To see this, recall that $\left\| D^{\theta}_{k(i,j)} \right\| = \| f(\theta_k \mid \theta_i) - f(\theta_k \mid \theta_j) \| > \left\| \tilde{f}(\theta_k \mid \theta_i) - \tilde{f}(\theta_k \mid \theta_j) \right\| = \left\| \tilde{D}^{\theta}_{k(i,j)} \right\|$ if $f(. \mid .)$ exhibits stronger complementarities than $\tilde{f}(. \mid .)$. Therefore, when complementarities are stronger, the constraint on the lower productivity academic of a potential blocking pair (equation 11 if $t > 0$, equation 10 if $t < 0$) is relaxed, while the constraint on the higher productivity academic of a potential blocking pair (equation 10 if $t > 0$, equation 11 if $t < 0$) is tightened. It is the latter constraint however that is on average the binding constraint for a blocking pair[53], so that a tightening of this constraint makes it less likely that a blocking pair to the maximal positive assortative matching $\tilde{\mu}$ exists.

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

[53] To see this, note that $E\left[D^{\nu}_{k(i,j)}\right] = E\left[\nu_{k,i} - \nu_{k,j}\right] = E\left[\nu_{k,i}\right] - E\left[\nu_{k,j}\right] = 0$ by assumptions 1a and 1b, while $D^{\theta}_{k(i,j)} = f(\theta_k \mid \theta_i) - f(\theta_k \mid \theta_j) > 0$ for $i > j$.

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 matching if no stable matching exists. Through the blocking dynamics such a less than maximally positive assortative matching is reached with positive probability.

Proof of Proposition 3: I first show that, if α_i is larger post-reform, so that $\alpha_i^{post} > \alpha_i^{pre}$, blocking pairs to the maximal positive assortative matching $\tilde{\mu}$ exist with a smaller probability than in the pre-reform scenario. To see this, note that i and $i+t$, with $t>1$ form a blocking pair to the matching $\tilde{\mu}$ post-reform if:

$$\alpha_i^{pre} D^{\theta}_{i(i+1,i+t)} \leq \alpha_i^{post} D^{\theta}_{i(i+1,i+t)} \leq D^{\nu}_{i(i+1,i+t)} \tag{12}$$

and

$$\alpha_{i+t}^{post} D^{\theta}_{i+t(i,\tilde{\mu}(i+t))} \geq \alpha_{i+t}^{pre} D^{\theta}_{i+t(i,\tilde{\mu}(i+t))} \geq D^{\nu}_{i+t(i,\tilde{\mu}(i+t))} \tag{13}$$

Similarly, i and $i+t$, with $t<0$ form a blocking pair post-reform if:

$$\alpha_i^{post} D^{\theta}_{i(i+1,i+t)} \leq \alpha_i^{pre} D^{\theta}_{i(i+1,i+t)} \leq D^{\nu}_{i(i+1,i+t)} \tag{14}$$

and

$$\alpha_{i+t}^{pre} D^{\theta}_{i+t(\tilde{\mu}(i+t),i)} \leq \alpha_{i-t}^{post} D^{\theta}_{i+t(\tilde{\mu}(i+t),i)} \leq D^{\nu}_{i+t(\tilde{\mu}(i+t),i)} \tag{15}$$

Post-reform, the constraint on the lower productivity academic of a potential blocking pair (equation 13 or 14) is relaxed, while the constraint on the higher productivity academic of a potential blocking pair (equation 12 or 15) is tightened. It is the latter constraint however that is on average the binding constraint for a blocking pair[54], so that a tightening of this constraint makes it less likely that a blocking pair to the maximal positive assortative matching $\tilde{\mu}$ exists.

From the proof of Proposition 2 we know that stronger complementarities affect the blocking pair constraints in a similar way as a larger α_i; the constraint on the lower productivity academic of a potential blocking pair is relaxed, and the constraint on the higher productivity academic of a potential blocking pair is tightened because $\left\| D^{\theta}_{k(i,j)} \right\|$ is larger. Given that α_i multiplies $\left\| D^{\theta}_{k(i,j)} \right\|$, we have that $\alpha_i^{post} \left\| D^{\theta}_{k(i,j)} \right\| - \alpha_i^{pre} \left\| D^{\theta}_{k(i,j)} \right\| > \alpha_i^{post} \left\| \tilde{D}^{\theta}_{k(i,j)} \right\| - \alpha_i^{pre} \left\| \tilde{D}^{\theta}_{k(i,j)} \right\|$ if $f(. \mid .)$ exhibits stronger complementarities than $\tilde{f}(. \mid .)$. The difference in the extent to which the constraint on the lower productivity academic of a potential blocking pair is relaxed, and the constraint on the higher productivity academic of a potential blocking pair is tightened from pre- to post-reform is therefore larger if complementarities are stronger. It follows that the probability that a blocking pair exists decreases more from before to after the reform if complementarities are stronger.

Secondly, I need to show that any matching μ that was stable pre-reform and 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:

$$\alpha_i D^{\theta}_{i(i+2,j)} \leq D^{\nu}_{i(i+2,j)} \tag{16}$$

[54] To see this, note that $E\left[D^{\nu}_{k(i,j)}\right] = E\left[\nu_{k,i} - \nu_{k,j}\right] = E\left[\nu_{k,i}\right] - E\left[\nu_{k,j}\right] = 0$ by assumptions 1a and 1b, while $D^{\theta}_{k(i,j)} = f(\theta_k \mid \theta_i) - f(\theta_k \mid \theta_j) > 0$ for $i > j$.

and

$$\alpha_j D^{\theta}_{j(i,\hat{\mu}(j))} \geq D^{\nu}_{j(i,\hat{\mu}(j))} \tag{17}$$

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 (equation 16) is relaxed if $\alpha_i^{post} > \alpha_i^{pre}$. At the same time, $D^{\theta}_{j(i,k)} = D^{\theta}_{i+1(i,i+3)} > 0$, so that the post-reform constraint on the lower productivity academic of the potential blocking pair (equation 17) is relaxed too. 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 as used in the first part of the proof, it follows that the probability that $i, i+1$ is a blocking pair for $\hat{\mu}$ post-reform is larger if complementarities are stronger.

Appendix B

In this appendix, I describe each of the three input data sets separately, before providing a detailed description of the preparation, manipulation and matching procedures used to generate the eventual university panel data set used for the empirical tests in this paper. All data handling was done using Python, unless otherwise indicated.

Input Data sets

Kuerschners Deutscher Gelehrten Kalender (hereafter: DGK) is a bibliographic and bibliometric encyclopedia of academics affiliated with German, Austrian or Swiss universities. All people that 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 too. Academics that 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 that 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 13-07-2013), as well as a copy of the exports from this database taken at 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 that 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.[55]

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 notifications 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 to construct the individual and university level data sets for this study, so as to align with the years for which I have (activation) data from DGK.

The ISI Web of Science database (hereafter: ISI) is published 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.[56]

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.[57] As for habilitation or Lehrbefugnis obtainment, I extract the name and current title of the person concerned, the current affiliation 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 is taken 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.[58] Here too, I take the month and year of the FuL issue in which the announcement

[55] 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.

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

[57] 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).

[58] In case of multiple offers, accepted or appointed offers are always recorded first, followed by offers that are obtained. Rejected offers are recorded last. In case of only obtained offers, offers from German universities are recorded first, otherwise the order

was made to be the time when the qualification was obtained, backdated by four months to correct for the average printing lag.[59]

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[60], coding titles and positions in a unified way, and classifying a title or position as being tenured or non-tenured [61]. 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[62].

Matching up Databases

Academics appearing in FuL are matched with academics in DGK on the basis of their last name[63], subject area[64] and initials[65]. Furthermore, a potential match is discarded if:

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

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

c) a potential DGK match is reported 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[66]

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

Publications from ISI are matched 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[67]. Whenever a match is found, the publication count of a person in the given ISI publication year is augmented by 1. Academics that share the same last name, initials and field are discarded for the study,

is random.

[59] 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.

[60] 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.

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

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

[63] 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).

[64] 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).

[65] 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).

[66] Where we 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.

[67] (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)

to prevent attributing publications of multiple different academics to multiple academics sharing last name, initials and field.

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[68]; 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.[69][70] 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 offer university of the FuL "erhalten" announcement is matched by the university recorded in DGK, this affiliation is filled forward as above.[71] 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.

From Individual-Level to University-Level Panel

For each academic, I derive 5 auxiliary variables. The 'start-date' of a person's academic career is taken the be 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. The 'end-date' of a

[68] This is based on self-reported information

[69] 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.

[70] 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.

[71] 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.

person's academic career is taken 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, the number of publication years, total publications and average number of publications in the 8 years up to and including year t, is calculated. In this time-span, only years in which the person is considered academically active are counted towards the number of publication years, and it is this number that is the denominator of the average number of publications. The mean and standard deviation of the latter three variables (number of publication years, sum of publications and average number of publications) are then calculated at the university level, over the people having a tenured affiliation at the university and not being retired yet. I use the year that someone turns 66 as the year in which he no longer takes his seat due to retirement [72]. I also calculate the number of people retiring from a given uni in a certain year, the number of new hires (defined as people not at the university x in year t-1, or at x but not in a tenured position), affiliates (people at the same uni, in a tenured position in year t-1) and the total faculty (sum of new hires and affiliates). The mean and standard deviation of the number of publication years, the sum of publications and average number of publications at the university are also calculated for new hires and affiliates separately. Furthermore, I calculate the difference between the mean of these variables for new hires and affiliates. Whenever the number of any group is 1 or less, the standard deviation of a variable is marked as missing, and similarly for the mean and difference between new hires and affiliates whenever a group is non-existent in a given year. Finally, the rank of a university is derived by calculating the average of the mean of average number of publications of affiliates over all universities - not including missing values - and marking a university with a higher than average mean average number of affiliate publications as 'high' rank. The coefficient of deviation of any of the aforementioned variables is calculated by dividing the mean of a variable by its standard deviation, marking this variable as missing whenever the standard deviation (and mean) is.

[72] The legal retirement age in Germany is 65, so we take the year after this lustrum to be the year in which the pensioner's position may be refilled.