JYVÄSKYLÄ STUDIES IN BUSINESS AND ECONOMICS
179

Juho Jokinen

Essays on Wages, Promotions and Performance Evaluations

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BUSINESS AND ECONOMICS 179

Juho Jokinen

Essays on Wages, Promotions and Performance Evaluations

Esitetään Jyväskylän yliopiston kauppakorkeakoulun suostumuksella
julkisesti tarkastettavaksi yliopiston Vanhassa juhlasalissa (S212)
elokuun 18. päivänä 2017 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Jyväskylä University School of Business and Economics,
in building Seminarium, hall S212, on August 18, 2017 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Essays on Wages, Promotions and Performance Evaluations

JYVÄSKYLÄ STUDIES IN BUSINESS AND ECONOMICS 179

Juho Jokinen

Essays on Wages, Promotions and Performance Evaluations

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Editors
Tuomo Takala
Jyväskylä University School of Business and Economics
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

.

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7136-6

URN:ISBN:978-951-39-7136-6
ISBN 978-951-39-7136-6 (PDF)

ISBN 978-951-39-7135-9 (nid.)
ISSN 1457-1986

Jyväskylä University Printing House, Jyväskylä 2017

ABSTRACT

Jokinen, Juho
Essays on wages, promotions and performance evaluations
Jyväskylä: University of Jyväskylä, 2017, 146 p.
(Jyväskylä Studies in Business and Economics
ISSN 1457-1986; 179)
ISBN 978-951-39-7135-9 (nid.)
ISBN 978-951-39-7136-6 (PDF)

This dissertation comprises four empirical research articles that use personnel data from a large university and worker-level panel data from Finland to examine the determinants of wages, promotions and employee performance evaluations. The first article employs personnel data to examine the importance of worker output and job seniority as predictors of employee performance evaluations and promotions. The results suggest that better-performing employees – with output measured both in absolute terms and relative to peers – were more likely to be assigned higher performance grades and had a higher probability of being promoted to more complex jobs than their peers with similar characteristics but lower output. Additionally, the findings suggest that employees with more job seniority were evaluated as exhibiting higher performance than were their equally productive but less-experienced peers. The second article employs personnel data to evaluate the role of gender in internal promotion, employee performance evaluation and earnings determination. The results indicate that gender had no effect on the probability of being promoted, conditional on productivity, nor did it play a role in the performance evaluation of employees. Furthermore, the observed male premium in earnings was mainly attributable to individual differences in worker output and background attributes. The third article employs sample data to examine the responsiveness of the pay level to local unemployment conditions and to test the hypothesis that pay level is more responsive to the unemployment level in less agglomerated and more remote regions. The results consistently suggest that the pay level is lower in localities with a higher unemployment level. The results provide some evidence that the unemployment elasticity of pay varies across different regions of the country, but once the unobserved worker heterogeneity is controlled for, the elasticity is unrelated to the degree of regional agglomeration. The fourth article employs sample data to examine the implications of using alternative measures of unemployment to estimate the local unemployment elasticity of pay. The results illustrate that the estimate of the local unemployment elasticity of pay varies considerably depending on the unemployment measure and the estimation technique used in the analysis but is not sensitive to the level of geographical disaggregation at which unemployment is measured.

Keywords: performance evaluations, promotions, gender pay gap, wage curve

Author's address Juho Jokinen
School of Business and Economics
University of Jyväskylä
P.O.Box 35, FI-40014 University of Jyväskylä
juho.p.jokinen@jyu.fi

Supervisors Professor Jaakko Pehkonen
School of Business and Economics
University of Jyväskylä

Reviewers Tuomas Pekkarinen
VATT Institute for Economic Research
Aalto University, School of Business
Helsinki, Finland

Antti Kauhanen
Research Institute of the Finnish Economy (ETLA)
Helsinki, Finland

Opponent Tuomas Pekkarinen
VATT Institute for Economic Research
Aalto University, School of Business
Helsinki, Finland

ACKNOWLEDGEMENTS

I am deeply indebted to many people who helped me and supported me during this long project. First and foremost, I wish to express my sincere gratitude to my supervisor, Professor Jaakko Pehkonen, for his constant support and encouragement during my postgraduate studies. Jaakko co-authored two of the four research articles in this dissertation, and provided valuable ideas for the other two. Jaakko always had time to read my work and discuss any questions I had, and this dissertation would not have been possible without all his help and guidance. I also like to acknowledge Professor Kari Heimonen, who constantly spurs and encourages PhD students in their studies and is an important figure in creating the extremely pleasant working environment we have at the economics department. I also want to thank Mika Haapanen for his encouraging comments and answers to my various questions concerning estimation methods, data issues, practicalities etc.

Part of this thesis was written while I visited at the Aarhus University in Denmark. During my stay I was hosted by professor Tor Eriksson, to whom I would like to express my utmost gratitude for his warm hospitality.

I was very fortunate to have Antti Kauhanen and Tuomas Pekkarinen as reviewers of my thesis manuscript. I am thankful for their detailed comments and suggestions on how to improve the quality of the research articles included in this dissertation. It is a great privilege to have Tuomas Pekkarinen as my opponent.

This work would not have been possible without generous financial support from the Yrjö Jahnsson Foundation, the OP Group Research Foundation and the Finnish Cultural Foundation. Furthermore, Jyväskylä University School of Business and Economics provided funding and facilities that were crucial to the completion of this dissertation.

During this project, I have had the opportunity to meet many inspiring and delightful people. First I would like to acknowledge three former fellow-students who have become dear friends: Terhi Maczulskij, whose efficient working methods, skills and sense of humor I admire; Tuomo Suhonen, who is a great person to spend time with (and to travel with); and Hannu Karhunen, who is an awesome person in every sense of the word. When I started my postgraduate studies, you gave me valuable advice and a proper introduction to different aspects of academic life. We have had many enjoyable moments both in work-related contexts and in contexts that were not even remotely related to work (and hopefully have many more to come).

I also wish to acknowledge Matthias Strifler, whose brilliant insights and deep knowledge of different things constantly astonish me (and whose catchy laughter continually increases the well-being of the people around him). Furthermore, I want to acknowledge several people who joined our department along the way: Antti Sieppi, Jaana Kari, Jussi Heikkilä, Kristian Koerselman, Mika Nieminen, Samu Kärkkäinen, Susanne Syrén and Ville Seppälä. I have really enjoyed our thought-provoking discussions, especially during our half-past twelve lunch breaks. We have also had numerous memorable moments outside the workplace. I also thank Jari Hännikäinen for all the vivid conversations we have had since we met during the first-year doctoral courses.

Finally, I wish to thank my family (my parents Riitta and Jukka and my siblings Tiia and Joonas) for their continuous support throughout my life. In addition, I thank Ari, Juho, Timo, Tommi and Tuukka for just being great friends. Finally, and most importantly, I would like to thank my wife Essi; thank you for being there for me, none of this would have been possible without your consistent support and encouragement.

Jyväskylä, July 2017

Juho Jokinen

LIST OF ORIGINAL PUBLICATIONS

Article 1. pp. 38–66.
Jokinen, J. and Pehkonen, J. "Worker output and seniority in performance evaluations and promotions" (unpublished)

Article 2. pp. 67–101.
Jokinen, J. and Pehkonen, J. "Promotions and Earnings – Gender or Merit? Evidence from Longitudinal Personnel Data". *Journal of Labor Research* (forthcoming)

Article 3. pp. 102–122.
Jokinen, J. "The wage curve and local monopsony power" (unpublished)

Article 4. pp. 123–143.
Jokinen, J. "Alternative unemployment measures as predictors of regional pay differences" (unpublished)

CONTENTS

CHAPTER 1: INTRODUCTION

1 Background and motivation

The empirical literature demonstrates that employees who are very similar in many respects can earn substantially different wages. Theoretical research has identified several factors that can generate wage differentials among employees. One primary contributing factor is the worker productivity. Because employers can afford to pay higher wages for more productive employees, wages are expected to be positively related to worker productivity. In accordance with this assumption, the empirical research illustrates that the wage level generally increases with productivity-related worker attributes such as education level (Card 1999; Harmon et al. 2003), work experience and job seniority (Altonji and Williams 2005; Buchinsky et al. 2010). Relying on the human capital theory, the wage premiums of more educated, experienced and senior workers can be explained based on the acquisition of productivity-enhancing skills and knowledge by means of schooling, learning-by-doing and on-the-job training (Card 1999; Topel 1991; Altonji and Williams 2005)[1].

However, the empirical literature reveals that worker characteristics that are not directly attributable to worker productivity also create wage differentials between otherwise identical employees. For example, studies often find a wage penalty for workers who belong to ethnic minorities (e.g., Altonji and Blank 1999). Furthermore, female employees are typically found to earn less than their otherwise comparable male colleagues (Altonji and Blank 1999; Blau

[1] Furthermore, because senior employees possess more firm-specific human capital and other specific skills and knowledge, and are hence less easily replaced, an employer may pay higher wages for them in order to restrain them from leaving the firm. In addition, other explanations have been proposed in the literature for the positive wage effects of education, work experience and job seniority. For example, the education premium may be due to the fact that employers use education level as a signal for individual's inherent ability (e.g., Weiss 1995; Arcidiacono et al. 2010). The wage premium for job seniority may be attributable to employers' objective to provide incentives to employees or to reduce labor turnover (see, e.g., Ballou and Podgursky 2002).

and Kahn 2000). The lower wages of women and members of minority groups are often attributed to the discriminatory behaviour of employers. Moreover, some wage differentials may be a consequence of employers' favouritism towards some worker groups. For example, the seniority wage premium may be partly attributable to favouritism: supervisors may favour senior workers in compensation decisions, for example, because they may have closer interpersonal relationships with senior subordinates than with junior subordinates.

The empirical research further documents that wages may vary considerably even among employees who have identical background characteristics (including, for example, education level, job seniority and gender). These wage differences partly stem from differences in job attributes. For example, workers who are employed in more demanding and complex jobs generally earn more than those employed in simpler jobs. Consequently, because the degree of job complexity is typically higher at higher levels of organizational hierarchy, wage differentials partly arise from assignment of employees to different levels of the hierarchy.

The research literature also identifies a variety of factors that can cause wages to differ among employees who share identical background characteristics and are doing similar jobs. For instance, the characteristics of the region in which the workplace is located can play a role in wage determination. One such characteristic is the unemployment level of the region: empirical evidence suggests that wages are lower in regions with higher unemployment levels (Blanchflower and Oswald 1994, 2005; Nijkamp and Poot 2005). In addition, wages are generally found to be higher in more densely populated regions with higher concentrations of economic activity, as a number of studies have documented that employees in large cities earn more than identical employees in smaller cities and rural regions do (Glaeser and Maré 2001; Yankow 2006; Gould 2007). The theoretical literature provides a number of alternative explanations for these findings, one of which is the difference in employees' bargaining power. The bargaining power of an employee is likely to be an increasing function of his or her outside job opportunities (i.e., job chances in firms other than the current one). One important determinant of outside job opportunities is the employment conditions in the surrounding labour market. Owing to higher mobility costs (job search, commuting and migration costs), job opportunities in distant locations are less relevant for the employee, and hence, the outside options are largely determined by the job opportunities in the local labour market. The job opportunities in the locality, however, are partly contingent on the number of unemployed jobseekers: a rise in the local unemployment level increases the job competition and consequently reduces the outside job opportunities of the employees in the locality. Hence, a higher local unemployment level is likely to reduce the bargaining power of employees and cause negotiated wages to be lower in high-unemployment regions. Outside job opportunities are typically also poorer in regions with a lower degree of economic agglomeration, such as in small cities and urban regions, because of the smaller number of potential employers. When employees have only a few (or, in the extreme case, none) potential alternative employers in the locality, their employers may have monopsony power over them, which they may take advantage of to negotiate lower wages. The monopsony power of the employers is presumably particularly strong in low-agglomeration regions with high joblessness.

Differences in compensation schemes can also give rise to wage differentials. In particular, compensation schemes that link pay level with the employee's job performance will generate wage differentials among employees with different performance. Because the wages of the employees who are compensated through performance pay reflect differences in worker performance, they should be unrelated to worker characteristics that do not affect performance. This is likely to be the case when the pay level is explicitly tied to some objective performance measure(s), such as the number of units produced. However, because obtaining objective information on worker performance can be difficult, the determination of performance pay is often based on subjective evaluations of employee performance. Although subjective evaluations may potentially provide a more comprehensive view of job performance by accounting for work aspects that are not quantifiable or measurable, they are also prone to biases arising from supervisors' personal opinions, impressions and preferences. These subjective biases can disrupt the relationship between performance evaluations and actual worker performance, which can result in distortions of the within-firm wage structure when biased evaluations are used to make decisions about pay level and promotions.

This thesis provides an empirical analysis of the determinants of worker pay level and pay-related personnel decisions including subjective performance evaluations and promotions. Empirical estimations employ two worker-level data sets from Finland: a unique longitudinal personnel data set from of a large university and a 7% random sample of the Finnish population. The research articles of this thesis contribute to the empirical research literature in personnel economics, labour economics and regional economics by analysing two main research topics. First, the research articles that employ the personnel data analyse the role of worker output and worker characteristics in wage determination, subjective performance evaluations and promotion decisions. The main focus of these articles is on examining how these pay-related decisions depend on gender, job seniority and worker output (measured either in absolute terms or relative to peers). Second, the research articles that use the random sample examine the responsiveness of the pay level to local unemployment conditions. In particular, these articles analyse whether the pay level is more responsive to the unemployment level in less agglomerated and more remote regions (as might be expected if employers have a higher degree of local monopsony power in such regions) and test whether the estimates of the local pay-unemployment relationship are sensitive to the chosen estimation technique, unemployment measure and regional disaggregation.

The remainder of this introductory chapter is organized as follows: The next section provides an overview of the research literature that is related to the topics analysed in this thesis. Section 3 briefly describes the relevant institutional characteristics of the country being analysed (Finland). Section 4 provides an outline of the thesis.

2 Related research literature

2.1 Performance-related compensation schemes

Pay schemes that link worker compensation to job performance have become increasingly popular over the past decades and are currently used extensively to compensate workers employed in a wide variety of jobs in both private companies and in public organizations (Brown and Heywood 2002; Weibel et al. 2010; Lucifora 2015)[2]. The primary objective of performance-related pay schemes is to generate incentives for employees to exert greater work effort, which, in turn, is expected to translate into higher worker productivity. The empirical research literature provides strong supportive evidence for the beneficial productivity effects of pay-for-performance schemes relative to a non-contingent (fixed wage) scheme, at least in occupations where worker performance is easily observable and measurable (Lazear 2000; Shearer 2004)[3]. The empirical findings indicate that productivity gains of performance-based pay schemes are attributable to two factors. First, the adoption of these schemes increases, on average, the productivity of workers. Second, the implementation of pay-for-performance schemes may initiate a productivity-enhancing self-selection process: such schemes tend to attract highly skilled employees and to increase the retention of more productive workers and the attrition of low-performing workers (Bishop 1987; Dohmen and Falk 2011). In addition to the incentive motive, employers may have other motives to adopt pay-for-performance schemes. They may use these schemes, for example, to motivate employees to devote more working time to activities desired by the employer. Furthermore, employers may adopt these schemes to increase wage flexibility (e.g., Kauhanen and Piekkola 2006).

Employers can use alternative mechanisms to link compensation with worker performance, one of which involves tying compensation explicitly to observed measures of performance, such as units of output produced or some other accurate measure (e.g., personal sales volumes)[4]. Employing such explicit

---

2 In this section, we focus on pay schemes that relate compensation to worker-level performance. Alternatively, employers often use incentive schemes that tie compensation to performance measures that are assessed at a more aggregated level; for example, pay schemes that link compensation to team performance or firm profits are widely used (e.g., Long and Fang 2012; Bandiera et al. 2013).

3 However, as illustrated by Ariely et al. (2009), the productivity effects may depend on the level of performance pay, with very high reward levels leading to reduced performance.

4 This section covers only compensation schemes that create incentives to employees by linking rewards to work performance. However, theoretical studies also present

contracting, however, requires employers to have detailed, objective information on their employees' work output. The extensive theoretical literature on the optimal responsiveness of compensation to worker performance highlights several factors that may hamper the collection of accurate information on worker performance (e.g., Holmström 1979; Lazear and Rosen 1981; Holmström and Milgrom 1991; Baker et al. 1994): monitoring work effort can be costly, some aspects of work effort are not directly observable or measurable (e.g., social interaction), and the available performance measures can be inaccurate (e.g., due to measurement error or manipulation by employees) and can vary for reasons unrelated to work effort (e.g., lower sales during an economic downturn).

The absence or noisiness of worker-specific performance measures make it more difficult for employers to identify those employees who currently outperform their peers – and who should thus be rewarded with wage bonuses or promotions to higher paying jobs. However, even when accurate information on worker performance is available, employers may be reluctant to tie compensation to specific performance measures because doing so may lead to undesirable behavioural effects: linking compensation to specific observable measures of job performance may cause employees to focus their work effort on these rewarded activities at the expense of other, non-rewarded activities (Holmström and Milgrom 1991; Baker 1992). Furthermore, tying compensation to quantifiable performance measures may induce employees to strive for a higher output at the expense of output quality (Paarsch and Shearer 2000). Due to these potentially adverse behavioural effects, together with the inaccuracies and limitations of observable performance measures, employers may base their compensation decisions on subjective employee performance evaluations. Thus, supervisors may develop subjective assessments of the job performance of their subordinates (possibly by incorporating available information about actual work effort into their assessments), which are then used as a basis for rewards for better-performing employees (e.g., Prendergast and Topel 1993). Alternatively, employers may find it optimal to adopt contracts that tie employee compensation to combined information on observed performance measures and subjective performance appraisals (Baker et al. 1994).

Subjective performance appraisal introduces an element of supervisory discretion and has the advantage of accounting for work aspects that are not easily quantified and measured, thus enabling a comprehensive evaluation of job performance. However, subjective appraisal also leaves more room for a supervisor's personal opinions, impressions and preferences (Bol 2008, 2011; Bol and Smith 2011), which may, in the absence of appropriate incentives for supervisors, lead to biased evaluation outcomes. The empirical literature re-

several alternative mechanisms that employers can use to provide incentives for a greater work effort. For example, the promotion of relatively better performers to higher paying jobs may serve as an incentive for employees to increase their work effort (Lazear and Rosen 1981). Alternatively, dismissal of underperforming employees may act as an incentive for work effort (Shapiro and Stiglitz 1984). For comprehensive reviews of theoretical research on performance pay schemes and other incentive schemes, see Lazear (1995), Prendergast (1999) and Lazear and Oyer (2013).

ports several biases that may cause subjective appraisals to be inconsistent with employees' actual work performance. First, such appraisals may be affected by worker characteristics unrelated to productivity, such as race (Elvira and Town 2001) and gender (Bartol 1999; Castilla 2012), and supervisors appear to rate subordinates with attributes similar to their own more favourably. For example, evaluators tend to assign higher ratings to subordinates of their own race (Kraiger and Ford 1985; Elvira and Town 2001). Second, supervisors may be reluctant to differentiate among employees, causing performance ratings to cluster around certain "norm" ratings (e.g., Murphy 1992). Consequently, performance ratings are less variable than would be expected based on the variance in actual work performance. Third, supervisors tend to show leniency in performance appraisals by rating their subordinates higher than the employees' actual performance warrants, particularly when the ratings are used for administrative purposes, such as determining wage increases and promotions (e.g., Jawahar and Williams 1997; Moers 2005).

Supervisors may have different motives for giving biased performance appraisals, including a reluctance to communicate poor evaluations to employees, a desire to curb wage increases when performance ratings are attached to wage levels and having the inclination to favour some employees over others (e.g., Prendergast & Topel 1993). Subjective biases in evaluations can disrupt the relationship between performance ratings and actual worker performance, which may have many undesirable consequences (e.g., Prendergast and Topel 1993; Moers 2005; Golman and Bhatia 2012). Most notably, such biases reduce the informational value of performance evaluations and can result in inaccurate personnel decisions when biased evaluations are used to make decisions about pay raises, promotions and dismissals. The misallocation of pay raises and unwarranted promotions to higher paying jobs may, in turn, result in distortions of the within-firm wage structure. Such distortions will reduce incentives for both poor and good performers and may lead to lower productivity, higher labour turnover costs and the loss of competent workers with extensive, firm-specific human capital. Additionally, biases in subjective performance ratings reduce the usability of these ratings for research purposes, as economic studies often use subjective performance ratings as a proxy for actual worker productivity (e.g., Flabbi and Ichino 2001; Pema and Mehay 2010).

One prominent feature of pay-for-performance schemes is that they should eliminate wage differentials that are not related to job performance: workers with equal work effort or output should earn similar salaries irrespective of their background characteristics. Consequently, performance pay schemes are likely to reduce the importance of worker characteristics (e.g., gender and race) and other factors (e.g., local unemployment conditions) in the remuneration of employees[5]. However, this may not hold in practice because,

5 There is some evidence to suggest that worker characteristics that are unrelated to productivity play no role in pay determination when compensation is based on job performance. For example, Heywood and O'Halloran (2005) found that racial wage differentials existed only for employees who were paid time rates and not for employees who were paid output-based pay.

with an exception of piece rate pay schemes, total salary is often only partly tied to job performance; hence, the fixed component of salary is prone to influences from irrelevant factors. Furthermore, performance rewards may differ across worker groups (owing, for example, to the higher bargaining power of some worker groups) and performance pay is liable to supervisory discretion when it is determined based on the outcome of subjective performance evaluation rather than some objective measure(s) of job performance.

While there exists an extensive body of theoretical work on various aspects of performance-related compensation schemes, the related empirical research remains limited due to a lack of suitable data sets that include information on worker performance measures. In the absence of more direct performance measures, prior studies on wages and promotions have often used subjective performance evaluations (e.g., Medoff and Abraham 1980, 1981; Flabbi and Ichino 2001; Dohmen 2004; DeVaro 2006; DeVaro and Waldman 2012) or performance-related pay components (e.g., Cassidy et al. 2016; DeVaro and Kauhanen 2016) to account for the differences in worker performance. These studies generally show that differences in worker performance translate into differences in rewards, indicating that employers use information on job performance to award rewards: better-performing employees are typically found to have higher wages and a higher probability of being promoted than their less well-performing colleagues with equal qualifications[6]. Furthermore, data from the academic labour market have provided an ideal setting for an analysis of the relationships between employee rewards and job performance, because the data on academic employees frequently include detailed individual performance measures, such as research productivity and teaching merit. The existing empirical evidence suggests that compensation schemes in academia reward good performance: a large body of evidence illustrates that increased academic performance is related to higher pay levels (e.g., Moore et al. 1998; Bratsberg et al. 2010) and higher promotion rates (e.g., Ginther and Hayes 2003; Haeck and Verboven 2012).

2.2 Gender differences in compensation

A vast body of empirical literature indicates that a gender wage gap is a persistent feature of labour markets (e.g., Arulampalam et al. 2007). The findings of the empirical research illustrate that the gap in average wages of male and female workers ("unadjusted gender wage gap") largely reflects gender differences in background and job characteristics. Most notably, men and women tend to segregate to different occupations and industries (Blau and Kahn 2000). Taken together with the fact that female-dominated occupations and industries typically have lower wage levels than male-dominated occupations and industries (e.g., Macpherson and Hirsch 1995), the gender segregation will result in

---

6 A recent study by Frederiksen et al. (2017) combines personnel data from several previously published studies to provide consistent evidence that subjective evaluations of work performance are positively related to pay levels and promotions.

gender pay gap of otherwise identical male and female workers[7]. Furthermore, the empirical findings suggest that the unadjusted gender pay gap is partly attributable to differences in productivity-related worker attributes. For example, female workers typically have more career breaks due to childbearing and heavier family and household responsibilities and, hence, have less job seniority and work experience (e.g., Waldfogel 1998). The higher frequency of career breaks will interrupt career advancement and result in a lower pay level, for example, due to lower seniority-related wage bonuses. Finally, female workers' lower probability of achieving the high-paid jobs of the organizational hierarchy provides yet another explanation for the gender wage gap, as empirical findings suggest that women are less likely to be promoted than are equally qualified men (e.g., Pergamit and Veum 1999; Pekkarinen and Vartiainen 2006).

However, although many countries have enacted equal pay legislation to enforce employers to pay equal wages for male and female employees with similar qualifications and work experience (e.g., Polachek 2014), the research literature reveals that women earn less than men do, even after accounting for differences in background and job characteristics (e.g., Altonji and Blank 1999)[8]. The remaining gender wage gap that is unexplained by differences in worker and job characteristics is often attributed to discriminatory behaviour towards female workers. Theoretical studies propose alternative motives for the gender-based pay discrimination. The gender gap in pay may be attributable to employers' discriminatory tastes against female workers (Becker 1957)[9]. Alternatively, the male-female pay gap may result from 'statistical discrimination' by employers. According to the statistical discrimination models, employers may use information on gender-specific averages of productivity-related characteristics to form expectations about the productivity of individual employees and pay lower wages for female workers when this information leads them to expect that women are less productive (e.g., Phelps 1972; Aigner and Cain 1977). Employers' monopsony power provides yet another explanation for the pay discrimination against female employees: given that women have a lower elasticity of labour supply than men do, monopsonistic employers may pay lower wages for female employees than male employees without inducing an increase in the quit rate of women (Barth and Dale-Olsen 2009; Hirsch et al. 2010).

However, the research literature also provides supportive evidence for a set of alternative explanations for the remaining unexplained gender pay gap. For example, female employees may be less willing to apply for pay raises and

---

7 The differences in pay levels of occupations/industries with different shares of female workers imply that the gender pay gap is partly attributable to gender differences in educational decisions made prior to entering the labor market: women may more often attain educational degrees that qualify for a job in female-dominated occupations (e.g., Blau and Kahn 2000).

8 In fact, some studies have illustrated that enforcement of equal pay legislation does not necessarily lead to the narrowing of the gender pay gap (e.g., Shapiro and Stelcner 1987).

9 According to Becker's (1957) model of the taste-based discrimination, gender-biased wage outcomes may also result from co-workers' or customers' discriminatory tastes against female employees.

promotions (Babcock and Laschever 2003; Booth 2009). The gender pay gap may also reflect gender differences in outside job opportunities: female employees may receive fewer outside job offers than do their male colleagues, which may reduce their bargaining power in the wage negotiations (Blackaby et al. 2005). Furthermore, gender differences in preferences and allocation of working time between job duties with different rewards may explain the male wage premium (e.g., Daymont and Andrisani 1984). Finally, gender differences in attitudes towards risk, competition and negotiation may contribute to the gender gap in wages (Bertrand 2011).

The unexplained gender gap in wages may also reflect gender differences in productivity (job performance). In the absence of information on worker productivity, studies on gender pay gap often add productivity-related proxy variables, such as job seniority and educational level, to control for productivity differences. These proxy variables may only partly account for actual differences in worker productivity, and hence, the observed gender pay gap may suffer from omitted variable bias. However, the empirical evidence illustrates that gender differences in compensation also exist in labour markets in which the worker productivity is more easily observed and measured. One such labour market is that of university faculty. The research on gender differences in labour market outcomes in academia illustrates that female faculty members earn less (e.g., Toutkoushian 1998; Blackaby et al. 2005) and are less likely to be promoted (e.g., Ward 2001; Ginther and Hayes 2003) than are male colleagues of comparable merit and productivity.

Based on the discussion above, it may be misleading to conclude that the unexplained gender wage gap (the gap that is observed after adjusting for differences in background characteristics) is entirely attributable to discriminatory behaviour. The reason is that the estimate of the gender wage gap may suffer from omitted variable bias: estimated wage regression may exclude regressors that partly determine wage level and correlate with gender, causing the gender pay gap estimate to be biased. However, while it may be tempting to conclude that there exists no gender-based discrimination in wage-setting when wage regression yields a negligible coefficient estimate on the gender variable, this conclusion can be misleading if one or more of the included control variables are partly determined by gender. For example, if female workers are less likely to achieve occupations higher up the organizational hierarchy than otherwise comparable male workers (causing the occupation variable to be gender-biased), a wage regression that controls for occupation may yield an underestimate of the true gender wage gap (see, e.g., Neal and Johnson 1996; Becker and Toutkoushian 2003)[10].

Empirical findings demonstrate that the gender gap in wages has narrowed over time in many countries (e.g., Blau and Kahn 2000). The reduction in

---

10 While drawing definite conclusions about the gender-based discrimination from observational data can be challenging, studies that have employed experimental techniques to assess discrimination in labor market outcomes (in particular in hiring) often provide strong evidence to suggest that women are discriminated against in the labor market (see, e.g., Azmat and Petrongolo 2014).

the gender gap is attributable to various changes in the labour markets. In particular, the wage inequality between genders has been reduced as the increased labour market participation and educational attainment of women have enabled them to enter better-paid occupations and industries that were previously occupied by men (Blau and Kahn 2000; Mulligan and Rubinstein 2008). Moreover, the increase in the incidence of performance-based pay schemes may potentially have promoted gender pay equality. Based on the discussion in the previous section, the gender pay gap may be lower for employees compensated by means of performance-based pay schemes, because in these schemes, equally productive employees should earn similar pay, regardless of their background characteristics. However, the empirical research on whether this is in fact the case is inconclusive: some findings suggest that the gender pay gap is smaller when workers are paid on the basis of output rather than on the time they spent working (Jirjahn and Stephan 2004; Petersen et al. 2007), while others indicate that gap is more pronounced in pay-for-performance wage systems (De la Rica et al. 2010; Kangasniemi and Kauhanen 2013). This finding may be explained by the fact that employers may be able to alter the rewards for better job performance between male and female workers. Moreover, when performance pay is based on a subjective evaluation of job performance by employee's supervisor, the gender pay gap may arise from gender differences in evaluation outcomes (Bartol 1999; Castilla 2012). Furthermore, the total salary of the employees is typically only partly dependent on personal job performance, and the fixed component of salary may be prone to gender discrimination.

2.3 Responsiveness of wages to local unemployment conditions

The exhaustive empirical work of Blanchflower and Oswald (1990, 1994) on the responsiveness of wages to local unemployment conditions in twelve countries yielded two notable conclusions. First, their findings from the analysis of worker-level data sets consistently showed that the wage level was negatively related to the local unemployment level, indicating that wages were lower in regions with higher unemployment. This finding contradicted the paradigm of the compensating wage differential theory, which predicts a positive relationship between local unemployment and the pay level (Harris and Todaro 1970). According to this theory, employees of the high-unemployment regions earn compensating wage differentials for the higher unemployment risk in these regions, and therefore, wages are positively related to the local unemployment level. Blanchflower and Oswald labelled the negatively sloping local pay-unemployment relationship as the "wage curve". Subsequent empirical support for the wage curve relationship is overwhelming, as it has been confirmed in over 40 countries (Nijkamp and Poot 2005; Blanchflower and Oswald 2005)[11].

---

[11] Blanchflower and Oswald (1994) proposed efficiency wages as one possible explanation for the wage curve relationship. According to the efficiency wage theory (e.g., Shapiro and Stiglitz 1984), an inverse wage-unemployment relationship arises because the higher local unemployment worsens the outside options of the employees, which reduces the wage levels needed to deter workers from shirking. In addition to

Second, their results suggested that the magnitude of the regional pay-unemployment relationship, which is commonly measured using the unemployment elasticity of pay, was substantially similar in all twelve countries analysed: the findings consistently indicated that a ten-percent increase in the local unemployment rate resulted in a one-percent decrease in wage level. Based on this empirical observation, Blanchflower and Oswald concluded that the responsiveness of wages to local unemployment did not seem to be affected by cross-country differences in the labour market institutions of countries. The magnitude of the wage curve relationship is of considerable economic interest because it measures the degree of wage flexibility with respect to local unemployment conditions. In contrast to the findings of the pioneering work of Blanchflower and Oswald (1990, 1994), subsequent studies reveal that this magnitude varies considerably across countries (see, e.g., the meta-analysis of Nijkamp and Poot 2005). The findings of the later studies also indicate that local wage responsiveness is contingent on labour market institutions. Indirect evidence is based on less elastic wage curves in countries with more centralized wage bargaining systems, such as those in Nordic countries (see, e.g., Albæk et al. 2000; Nijkamp and Poot 2005), and union workers (e.g., Card 1995; Barth et al. 2002), and more direct evidence is provided by studies illustrating that the slope of the wage curve changes as a result of the restructuring of the wage bargaining system and other labour market reforms (Devicienti et al. 2008; Cholezas and Kanellopoulos 2015; Daouli et al. 2017).

3 Institutional background

3.1 Wage setting in Finland

Wage setting in Finland is characterized by a high centralization of bargaining, with wage bargaining occurring between the labour unions and employer organizations[12]. Wage negotiations are typically conducted either at the sectoral level or at the level of the federations of labour unions and employer organizations. In the latter case, the government has often participated in wage negotiations, in which case the negotiated collective agreements are typically joined with changes in taxation, labour market policies and pensions, social security and unemployment insurance systems (the resulting agreements of such tripartite negotiations are referred to as comprehensive incomes policy agreements).

efficiency wage theory, Blanchflower and Oswald relied on the labor contract model and union bargaining model to develop alternative theoretical explanations for the wage premium of the workers in low-unemployment regions. Later studies by Campbell and Orszag (1998) and Sato (2000) introduce additional theoretical explanations for the wage curve relationship, building upon an extension of the efficiency wage model and a search theory, respectively.

12 There is a long tradition of highly centralized (and coordinated) wage bargaining in Finland: the key features of the wage bargaining system have been substantially similar since the late 1960s (Uusitalo and Vartiainen 2009).

Negotiated sectoral collective agreements specify the minimum pay levels of different occupations, general wage increases and other minimum terms of employment.

Owing to the high union membership rate – in 2015, 76% of the employees were members of labour unions (Statistics Finland 2016) – the coverage of the collectively bargained wage agreements is very high. All public sector employees are covered by the collectively bargained agreements. Moreover, when the sectoral collective agreements are declared generally applicable, these agreements extend to cover also many of the non-unionized private sector employees. The general applicability of the sectoral collective bargaining agreements is determined by a special commission, which is appointed by the government and guided by legislation (Act on Confirmation of the General Applicability of Collective Agreements). In a typical case, a sectoral collective agreement is declared generally applicable when half of the employees in the sector work in firms that are members of an employers' association. In 2014, approximately 76% of private sector employees worked in firms that were members of employer organizations; however, due to generally applicable collective agreements, over 84% of all private sector employees were covered by the collective agreements. For employees both in the private and public sectors, the share was approximately 89% (Ahtiainen 2016). The wage formation of the employees not covered by the collective agreements (in 2014, these employees comprised approximately 11% of all employees) is subject to few restrictions. Their wage formation is not restricted by the occupation-specific minimum wage levels and general wage increases specified in the collective agreements. Furthermore, there is no statutory minimum wage in Finland. However, the legislation states that the employees not covered by the collective agreements (or other wage contracts) will be paid a 'customary and reasonable' wage (Employment Contracts Act).

Following the international trends (e.g., Brown and Heywood 2002), various types of performance-related pay schemes have become more common practice for worker compensation also in Finland over the last few decades (Piekkola 2005; Kauhanen and Piekkola 2006). The increased implementation of such pay schemes may reflect employers' objective to increase flexibility in wage setting[13]. In the private sector, for example, manufacturing industries have widely adopted pay schemes that link employer compensation (partly) with job performance. The performance-based pay schemes used in the manufacturing industries include, for example, piece-rate schemes, team bonus schemes and schemes that tie total earnings partly to employee performance evaluations (Pekkarinen and Vartiainen 2006; Pekkarinen and Riddell 2008; Pekkarinen and Uusitalo 2012).

In addition to private sector companies, pay schemes that tie remuneration partly to job performance have also been widely adopted in the public sector. The compensation of employees in the central government and state-

[13] In fact, survey responses of Finnish manufacturing employers reveal that increasing wage flexibility has been one of the objectives in the adoption of performance-related pay schemes (see Kauhanen and Piekkola 2006).

funded organizations is generally based on a salary system that determines total earnings based on two factors: the complexity of the job duties and personal job performance. In this salary system, the base salary level of an employee is determined by his or her position on a job complexity ladder, where each rung of the ladder is attached to a predetermined base salary level. The base salaries attached to complexity rungs are established by a collective bargaining agreement. In addition to complexity level-specific based salaries, the salary system uses subjective performance ratings to determine salary variation within the complexity levels; each performance grade increases the complexity level-specific based salary by a predetermined percentage[14].

The centralized wage setting and generally binding collective agreements may reduce the responsiveness of wages to factors such as firm-specific demand conditions and local labour market conditions. Importantly for the analysis of this thesis, wages may be less responsive to the unemployment level in local labour markets of Finland than in countries where wage bargaining is conducted at a more decentralized level[15]. In other words, the slope of the Finnish wage curve relationship may be smaller than those in countries with more flexible wage-setting systems. In fact, consistent with this assumption, Albæk et al. (2000) found no evidence of a negative relationship between wages and the regional unemployment rate in Finland. However, contrary to the findings of Albæk and others, the results in Pekkarinen (2001) and Maczulskij (2013) suggest that wages respond to local unemployment conditions also in Finland, hence confirming the existence of the wage curve relationship. Their findings underline the fact that although the collective bargaining agreements limit the downward flexibility of wages by defining the general wage increases and the minimum wage levels of different occupations, the Finnish wage-setting system still leaves plenty of room for local wage flexibility[16]. For example, employers can pay wages (wage increases) that exceed contractual wages (wage increases). Furthermore, based on mutual agreement between employers and employees, locally negotiated wage increases can deviate from general wage increases. These features can both potentially increase the responsiveness of wages to local labour market conditions.

---

14 A very similar pay scheme is also employed in some private sector manufacturing industries (see Pekkarinen and Vartiainen 2006).

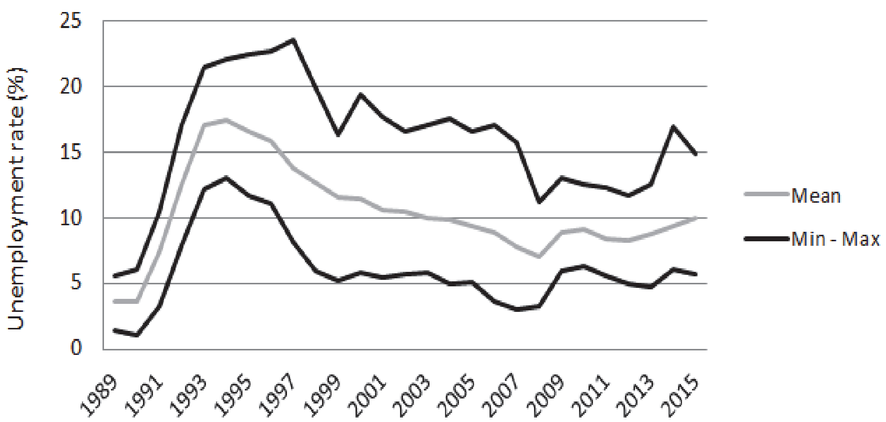
15 The findings of Blien et al. (2013) provide some evidence to suggest that wages are more responsive to local unemployment conditions (i.e., the slope of the wage curve relationship is larger) when wage bargaining occurs at a more decentralized level.

16 For a detailed description of the Finnish wage-setting system, see Uusitalo and Vartiainen (2009).

3.2 Regional unemployment in Finland

In the following empirical analysis of the responsiveness of the pay level to local unemployment conditions, both NUTS-3 regions (Nomenclature of Units for Territorial Statistics) and LAU-1 regions (Local Administrative Units) are used as local labour markets[17]. The NUTS-3 classification disaggregates mainland Finland into 19 regions and the LAU-1 classification into 79 regions. Local unemployment conditions vary substantially across different regions of Finland. Figure 1 reports the changes in the unemployment rates of the NUTS-3 regions between 1989 and 2015. At the beginning of the 1990s, the overall unemployment level in Finland was very low; in 1990, unemployment rates for the NUTS-3 regions varied between 1% and 6%. In 1991–1993, Finland underwent a serious economic recession (often referred to as the Finnish Great Depression), which led to a sharp increase in the regional unemployment rates, with some regions suffering more than others: in 1994, the unemployment rates for the NUTS-3 regions ranged from 13% to 22%. The recession was followed by a long period of uninterrupted economic growth (in 1994–2007, the average growth rate of the GDP was 4%), which led to a gradual decrease in overall unemployment. However, the notable regional differences in unemployment have proven to be very persistent: in 2015, the unemployment rates for the NUTS-3 regions varied between 6% and 15%.

FIGURE 1 Unemployment in NUTS-3 regions (1989–2015)



Notes: Mean, maximum and minimum unemployment rate of the NUTS-3 regions (before 2011, 19 regions; after 2010, 18 regions; Åland Islands excluded). Unemployment figures are based on the Labour Force Survey (Statistics Finland).

17 The borders of the LAU-1 sub-regions are in part determined to minimize the cross border commuting flows and are typically considered reasonable approximations of local labor markets.

4 Outline of the thesis

This thesis comprises four empirical research articles. Based on the research topics, the thesis can be divided into two parts. The first part of the thesis analyses the determinants of wages, promotions and employee performance evaluations and provides answers to the following research questions:

- Do subjective performance evaluations of employees reflect actual differences in worker output? Are better-performing workers more likely to be rewarded with higher performance evaluations and promotions than their equally qualified co-workers with poorer work performance? Are senior workers more likely to be rewarded with higher performance grades and promotions than their less experienced colleagues with equal qualifications and work output? (Chapter 2)

- Are female workers less likely to be rewarded with better performance evaluations and promotions than their male colleagues with equal qualifications and work output? After accounting for differences in background characteristics and work output, do women earn less than men? (Chapter 3)

The second part of the thesis examines the responsiveness of pay level to local unemployment conditions and seeks to address the following research questions:

- Is there a relationship between the wage level and the local unemployment rate? Does the responsiveness of pay level to local unemployment conditions depend on the degree of regional economic agglomeration? (Chapter 4)

- Is the slope estimate of the local pay-unemployment relationship sensitive to the choice of the unemployment measure or estimation technique? (Chapter 5)

Chapters 2 and 3 provide econometric case studies that utilize a longitudinal personnel data set from a large Finnish university to analyse the role of worker attributes and work output in employee performance evaluations, promotions and wage determination. Although results based on a single organization should be interpreted with some caution, there are many advantages of using such data to study different career outcomes. In particular, the data from personnel records allow one to analyse compensation and careers within an internal labour market with homogenous personnel policies and uniform criteria for remuneration and career advancement. Furthermore, personnel data are typically highly accurate and contain detailed information on attributes not available in customary survey and administrative data sets.

The personnel data used in this thesis provide unique information on worker performance and job hierarchy that allows us to address some of the limitations of previous studies of the same topics. First, we use these data to examine how subjective evaluations of worker performance depend on different worker attributes when differences in actual work output are accounted for. Additionally, the data set provides a detailed complexity ladder of job duties that enables us to overcome challenges related to the determination of job hierarchy, allowing us to analyse promotions along a well-defined job ladder. Furthermore, the data allow us to analyse whether promotion and pay decisions are based on objective and subjective measures of worker performance when accounting for differences in worker background characteristics.

The empirical analyses of Chapters 2 and 3 provide insights into compensation decisions under a formal salary system in which a subjective evaluation of job complexity and personal work performance determine an employee's total salary level. The structure of the university's pay scheme provides a useful case for analysing the roles of personal job performance and worker attributes in total compensation. In this pay scheme, each employee is assigned to one of the eleven rungs of the job complexity ladder, which determine the base salary levels, and to one of the nine performance grades, which determine the wage range within each complexity rung. Compensation is not explicitly linked to observable performance measures or worker attributes; hence, worker out-put and attributes can influence salaries only indirectly by affecting supervisors' subjective judgements regarding whether employees should be awarded with higher performance grades or be promoted to higher rungs of the complexity ladder.

Chapter 2 examines the extent to which differences in subjective performance evaluations can be attributed to differences in work output and job seniority. In addition, it analyses the effects of work output, job seniority and subjective performance evaluations on the probability of promotion along the job ladder. A descriptive analysis of the personnel data leads to several interesting conclusions regarding employee performance evaluations and promotions. First, the mean performance grade increased along the job ladder, indicating that employees higher in the job hierarchy also tended to receive higher performance grades. Second, the analysis illustrates that within-job level performance grades tended to cluster into a few "norm" grades. Third, the findings suggest that there was notable downward rigidity in the subjective performance evaluations, as performance grades were almost never downgraded. Fourth, the descriptive analysis reveals that demotions along the job complexity ladder were rare.

The analysis of the determinants of performance evaluations reveals that when comparing employees with equally complex job duties and similar background characteristics, the employees who outperformed their peers (i.e., those who had a higher relative work output) were more likely to have above-average performance grades. Moreover, employees with higher (relative) work output tended to have higher performance grades than their equally qualified colleagues with lower output. Additionally, the findings indicate that better-performing employees – with output measured both in absolute terms and relative to peers – had a higher probability of being upgraded to higher performance grades and being promoted to more complex jobs than their peers with similar characteristics but lower output. The analysis of performance evaluations also reveals that employees with more job seniority (measured in years of university service) tended to have higher performance grades than their less experienced peers with equal qualifications and output.

The promotion analysis illustrates that employees with higher performance grades had a higher probability of promotion, suggesting that supervisors also used prior evaluations of employee performance to guide their promotion decisions. Furthermore, the findings provide some evidence to suggest that more senior employees had a higher probability of being upgraded to higher performance grades and being promoted, although the statistical significance of the job seniority estimates depends on the included control variables.

Chapter 3 evaluates the role of gender in earnings determination, performance evaluations and promotion decisions within well-defined job ladders while accounting for differences in background characteristics and worker output. The results suggest that male and female researchers were equally likely to be promoted, conditional on individual research productivity. An analysis of the determinants of employee performance evaluations revealed that gender played a negligible or no role in evaluation decisions. The observed male premium in earnings was mainly attributable to individual differences in research productivity and background characteristics: adjusting for these differences reduced the gender earnings gap from approximately 11% to approximately 1–2%. Moreover, once the full set of controls was included, the gender coefficient was no longer statistically significant. Additionally, the results demonstrate that female researchers had lower research output than did their male colleagues, even after conditioning on a set of worker and job characteristics. The results further suggest that female and male professors were paid equally, although female employees were less likely to work as full professors than equally qualified men were.

Additionally, the results indicate that higher research productivity was related to higher probabilities of being promoted to or working on the highest job ladders (academic ranks). The findings also confirm that more productive researchers received more favourable performance evaluations than others with similar background characteristics, implying that the available worker output information was effectively employed in the assessment of employee performance and was therefore likely to reduce the subjectivity of the evaluation process and result in more objective performance evaluations.

Chapters 4 and 5 employ longitudinal micro data from Finland to examine the responsiveness of pay level to local unemployment conditions. The micro-data analysed are based on a 7% random sample of the Finnish population drawn in 2001. The data from a sampling year were merged with data from preceding and subsequent years, and the resulting longitudinal data include information on the sampled individuals for the period from 1995 to 2006. The analyses in these chapters focus on non-agricultural private sector employees who lived in mainland Finland and, consequently, individuals who lived in the Åland Islands (which constitute an autonomous province of Finland) as well as individuals who were employed in the public sector or by the agriculture, forestry or fishing industries were excluded from the final sample. Additionally, employees aged under 18 and over 68 were excluded from the sample. For the wage curve analysis, the worker-level data were combined with regional data on local unemployment rates measured at the LAU-1 level (79 sub-regions) and at the NUTS-3 level (19 regions).

Chapter 4 examines the within-country variation in the local unemployment elasticity of pay. More precisely, the chapter examines whether the pay level is more responsive to the unemployment level in less agglomerated and more remote regions, as might be expected on the basis of the hypothesis proposed by Longhi et al. (2006). They argued that the within-country variation in the slope of the wage curve may arise from local monopsony power caused by regional differences in geographical remoteness and in the degree of economic agglomeration. Their reasoning is as follows: the combination of fewer jobs and higher job mobility costs (including job search, commuting and migration costs) in more remote low-agglomeration regions weakens the outside job opportunities of the workers in these regions; the poorer job opportunities give employers local monopsony power over their employees, and consequently, employers achieve greater flexibility in adjusting wages based on local unemployment conditions. Consequently, the wage curve relationship is more pronounced in remote low-agglomeration regions than in regions that have a high concentration of firms and are in close proximity to neighbouring regions.

The results consistently suggest that the pay level is lower in localities with a higher unemployment level and hence provide strong support for the so-called wage curve hypothesis, which predicts a negative relationship between local unemployment and the pay level. Furthermore, the results indicate that conditional on the local unemployment rate, the unemployment conditions in neighbouring regions do not play a role in determining the pay level. The results provide some evidence that the magnitude of the wage curve relationship, which is estimated based on the unemployment elasticity of pay, varies across different geographical areas of a country. Moreover, the findings indicate that once worker fixed effects are included to control for the composition bias resulting from the changing composition of the workforce (Solon et al. 1994), wage curve slopes are similar across regions with different degrees of economic agglomeration. Hence, the findings do not provide consistent support for the monopsony power hypothesis proposed by Longhi et al. (2006) and imply that the failure to control for the unobserved worker heterogeneity (composition bias) may explain why they found a more pronounced wage curve relationship for the low-agglomeration regions than for the high-agglomeration regions of

Western Germany. Further analysis based on a more direct measure of local monopsony power, namely, the number of own-industry establishments in the locality, supports a similar conclusion: the responsiveness of the pay level to local unemployment conditions is not stronger for workers whose employers potentially have more monopsony power over them.

Chapter 5 examines the sensitivity of the slope of the wage curve relationship to different measures (definitions) of unemployment and different levels of geographical aggregation at which the local unemployment is measured. In addition, the chapter analyses the responsiveness of the pay level to education level-specific unemployment conditions in the locality by estimating the wage curve relationship using regional unemployment rates disaggregated into four education levels. The empirical analysis consistently yields negative elasticity estimates, suggesting that the local unemployment level is negatively related to the pay level. However, the results illustrate that a statistically and economically significant inverse pay-unemployment relationship is typically detected only when worker fixed effects are controlled for. The estimate of the local unemployment elasticity of pay varies considerably depending on the unemployment measure used in the analysis but is not sensitive to the level of geographical disaggregation at which unemployment is measured. Modifying standard unemployment rates to exclude long-term unemployed and to include participants in active labour market programs has only a modest effect on elasticity estimates. Finally, the negatively sloping pay-unemployment relationship is detected only with overall regional unemployment rates, whereas education level-specific regional unemployment rates yield a statistically insignificant elasticity estimate close to zero.

Table 1 summarizes the research topics, data sets, empirical methodology and the main findings of the research articles in Chapters 2–5.

TABLE 1 An overview of the thesis

Ch.	Research topic	Empirical approach	Main findings
2	Worker output and job seniority as determinants of subjective performance evaluations and promotions	- Data: personnel data from a large Finnish university - Period: 2006–2012 - Methods: OLS, linear probability model, ordered probit	- Employees with higher (relative) work output and more job seniority had higher performance evaluations. - Better-performing employees had a higher probability of being upgraded to higher performance grades and being promoted.
3	Gender differences in subjective performance evaluations, promotions and earnings conditional on worker characteristics and worker output	- Data: personnel data from a large Finnish university - Period: 2006–2012 - Methods: OLS, linear probability model, ordered probit	- Gender has no effect on the probability of being promoted, conditional on research productivity. - Gender played a negligible or no role in the performance evaluation of employees. - The male premium in earnings was mainly attributable to individual differences in research productivity and background characteristics.
4	Variation in the slope estimate of the wage curve across regions with different degree of economic agglomeration	- Data: a 7% random sample of the Finnish population - Period: 1995–2006 - Methods: OLS, fixed effects regression, 2SLS	- Pay level is lower in localities with a higher unemployment level. - Conditional on the local unemployment rate, the unemployment conditions in neighbouring regions do not play a role in determining the pay level. - Once the unobserved worker heterogeneity is controlled for, the local unemployment elasticity of pay is unrelated to the degree of regional economic agglomeration.
5	The sensitivity of the slope estimate of the wage curve to alternative unemployment measures, regional disaggregations and estimation techniques	- Data: a 7% random sample of the Finnish population - Period: 1995–2006 - Methods: OLS, fixed effects regression, 2SLS	- The slope estimate of the wage curve varies considerably depending on the unemployment measure used. - The negatively sloping wage curve is detected only with overall regional unemployment rates, but not with education level-specific regional unemployment rates.

References

Ahtiainen, L. 2016. Työehtosopimusten kattavuus vuonna 2014 (in Finnish). *Ministry of Employment and Economy Publications*, Employment and entrepreneurship 11/2016.

Aigner, D.J. and Cain, G.G. 1977. Statistical theories of discrimination in labor markets. *ILR Review* 30(2), 175–187.

Albæk, K., Asplund, R., Barth, E., Blomskog, S., Guômundsson, B.R., Karlsson, V. and Strøjer Madsen, E. 2000. Dimensions of the wage-unemployment relationship in the Nordic countries: Wage flexibility without wage curves. In: Polachek, S. and K. Tatsiramos (ed.) *Research in Labor Economics*, Vol. 19. pp. 345–381.

Altonji, J.G. and Blank, R.M. 1999. Race and gender in the labor market. *Handbook of labor economics*, Vol. 3, Part C, pp. 3143–3259.

Altonji, J.G. and Williams, N. 2005. Do wages rise with job seniority? A reassessment. *ILR Review* 58(3), 370–397.

Arcidiacono, P., Bayer, P. and Hizmo, A. 2010. Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics* 2(4), 76–104.

Ariely, D., Gneezy, U. and Loewenstein, G. 2009. Large stakes and big mistakes. *Review of Economic Studies* 76(2), 451–469.

Arulampalam, W., Booth, A.L. and Bryan, M.L. 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review* 60(2): 163–86.

Azmat, G. and Petrongolo, B. 2014. Gender and the labor market: What have we learned from field and lab experiments?. *Labour Economics* 30, 32–40.

Babcock, L. and Laschever, S. 2003. *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press: Princeton, NJ, US.

Baker, G. 1992. Incentive contracts and performance measurement. *Journal of Political Economy* 100(3), 598–614.

Baker, G., Gibbons, R. and Murphy, K.J. 1994. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics* 109(4): 1125–1156.

Ballou, D. and Podgursky, M. 2002. Returns to seniority among public school teachers. *The Journal of Human Resources* 37(4), 892–912.

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.

Barth, E., Bratsberg, B., Naylor, R.A. and Raaum, O. 2002. Explaining variations in wage curves: Theory and evidence. University of Oslo, Department of Economics, *Memorandum No. 03/2002*.

Barth, E. and Dale-Olsen, H. 2009. Monopsonistic discrimination, worker turnover, and the gender wage gap. *Labour Economics* 16(5), 589–597.

Bartol, K.M. 1999. Gender influences on performance evaluations. In: G.N. Powell (Ed.) *Handbook of Gender and Work*. pp. 165–178. Thousand Oaks, CA, US: SAGE Publications.

Becker, G.S. 1957. *The Economics of Discrimination*. The University of Chicago Press: Chicago, IL, US.

Becker, W.E. and Toutkoushian, R.K. 2003. Measuring gender bias in the salaries of tenured faculty members. *New Directions for Institutional Research* 117, 5–20.

Bertrand, M. 2011. New perspectives on gender. *Handbook of Labor Economics*, Volume 4, Part B, 1543–1590.

Bishop, J. 1987. The recognition and reward of employee performance. *Journal of Labor Economics* 5(4), 36–56.

Blackaby, D., Booth, A.L. and Frank, J. 2005. Outside offers and the gender pay gap: empirical evidence from the UK academic labour market. *The Economic Journal* 115(501): F81–F107.

Blanchflower, D.G. and Oswald, A.J. 1990. The Wage curve. *Scandinavian Journal of Economics* 92(2), 237–242.

Blanchflower, D.G. and Oswald, A.J. 1994. *The Wage Curve*. The MIT Press: Cambridge, MA, US.

Blanchflower, D.G. and Oswald, A.J. 2005. The Wage curve reloaded. *NBER Working Paper* No. 11338.

Blau, F.D. and Kahn, L.M. 2000. Gender differences in pay. *Journal of Economic Perspectives* 14(4), 75–99.

Blien, U., Dauth, W., Schank, T. and Schnabel, C. 2013. The institutional context of an 'empirical law': The wage curve under different regimes of collective bargaining. *British Journal of Industrial Relations* 51(1), 59–79.

Bol, J.C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* 27(1): 1–24.

Bol, J.C. 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review* 86(5), 1549–1575.

Bol, J.C. and Smith, S.D. 2011. Spillover effects in subjective performance evaluation: bias and the asymmetric influence of controllability. *The Accounting Review* 86(4): 1213–30.

Booth, A.L. 2009. Gender and competition. *Labour Economics* 16(6), 599–606.

Bratsberg, B., Ragan, J.F. and Warren, J.T. 2010. Does raiding explain the negative returns to faculty seniority?. *Economic Inquiry* 48(3), 704–721.

Brown, M. and Heywood, J.S. 2002. *Paying for performance: an international comparison*. Armonk, NY, US: M. E. Sharpe.

Buchinsky, M., Fougère, D., Kramarz, F. and Tchernis, R. 2010. Interfirm mobility, wages and the returns to seniority and experience in the United States. *Review of Economic Studies* 77(3), 972–1001.

Campbell, C. and Orszag, J.M. 1998. A model of the wage curve. *Economics Letters* 59(1), 119–125.

Card, D. 1995. The wage curve: A review. *Journal of Economic Literature* 33(2), 785–799.

Card, D. 1999. The causal effect of education on earnings. *Handbook of Labor Economics*, Volume 3, Part A, 1801–1863.

Cassidy, H., DeVaro, J. and Kauhanen, A. 2016. Promotion signaling, gender, and turnover: New theory and evidence. *Journal of Economic Behavior & Organization* 126(A), 140–166.

Castilla, E.J. 2012. Gender, race, and the new (merit-based) employment relationship. *Industrial Relations* 51(S1): 528–562.

Cholezas, I. and Kanellopoulos, N.C. 2015. Labour market reforms in Greece and the wage curve. *Economics Letters* 136(1), 19–21.

Daouli, J., Demoussis, M., Giannakopoulos, N. and Laliotis, I. 2017. The wage curve before and during the Greek economic crisis. *Empirical Economics* 52(1), 59–77.

Daymont, T.N. and Andrisani, P.J. 1984. Job preferences, college major, and the gender gap in earnings. *The Journal of Human Resources* 19(3), 408–428.

De la Rica, S., Dolado, J.J. and Vegas, R. 2010. Performance pay and the gender wage gap: evidence from Spain. *IZA Discussion Paper Series* No. 5032.

DeVaro, J. 2006. Internal promotion competitions in firms. *The RAND Journal of Economics* 37(3), 521–542.

DeVaro, J. and Kauhanen, A. 2016. An "opposing responses" test of classic versus market-based promotion tournaments. *Journal of Labor Economics* 34(3), 747–779.

DeVaro, J. and Waldman, M. 2012. The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics* 30(1), 91–147.

Devicienti, F., Maida, A. and Pacelli, L. 2008. The resurrection of the Italian wage curve. *Economics Letters* 98(3), 335–341.

Dohmen, T.J. 2004. Performance, seniority, and wages: formal salary systems and individual earnings profiles. *Labour Economics* 11(6), 741–763.

Dohmen, T. and Falk, A. 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review* 101(2), 556–590.

Elvira, M. and Town, R. 2001. The effects of race and worker productivity on performance evaluations. *Industrial Relations* 40(4), 571–589.

Flabbi, L. and Ichino, A. 2001. Productivity, seniority and wages: New evidence from personnel data. *Labour Economics* 8(3), 359–387.

Frederiksen, A., Lange, F. and Kriechel, B. 2017. Subjective performance evaluations and employee careers. *Journal of Economic Behavior & Organization* 134, 408–429.

Ginther, D.K., and Hayes, K.J. 2003. Gender differences in salary and promotion for faculty in the humanities 1977–95. *The Journal of Human Resources* 38(1), 34–73.

Glaeser, E.L. and Maré, D.C. 2001. Cities and skills. *Journal of Labor Economics* 19(2), 316–342.

Golman, R. and Bhatia, S. 2012. Performance evaluation inflation and compression. *Accounting, Organizations and Society* 37(8), 534–543.

Gould, E.D. 2007. Cities, workers, and wages: A structural analysis of the urban wage premium. *The Review of Economic Studies* 74(2), 477–506.

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

Harmon, C., Oosterbeek, H. and Walker, I. 2003. The returns to education: Microeconomics. *Journal of Economic Surveys* 17(2), 115–156.

Harris, J.R. and Todaro, M.P. 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review* 60(1), 126–142.

Heywood, J.S. and O'Halloran, P.L. 2005. Racial earnings differentials and performance pay. *The Journal of Human Resources* 40(2), 435–452.

Hirsch, B., Schank, T. and Schnabel, C. 2010. Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from Germany. *Journal of Labor Economics* 28(2), 291 – 330.

Holmström, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10(1), 74–91.

Holmström, B. and Milgrom, P. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization* 7 (special issue): 24–52.

Jawahar, I.M., and Williams, C.R. 1997. Where all the children are above average: The performance appraisal purpose effect. *Personnel Psychology* 50(4), 905–925.

Jirjahn, U. and Stephan, G. 2004. Gender, piece rates and wages: evidence from matched employer-employee data. *Cambridge Journal of Economics* 28(5): 683–704.

Kangasniemi, M. and Kauhanen, A. 2013. Performance-related pay and gender wage differences. *Applied Economics* 45(36): 5131–143.

Kauhanen, A. and Piekkola, H. 2006. What makes performance-related pay schemes work? Finnish evidence. *Journal of Management and Governance* 10(2), 149–177.

Kraiger, K. and Ford, J.K. 1985. A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology* 70(1), 56–65.

Lazear, E.P. 1995. *Personnel Economics*. The MIT Press: Cambridge, MA, US.

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

Lazear, E.P. and Oyer, P. 2013. Personnel economics. In: Gibbons, R. and Roberts, J. (eds.) *The Handbook of Organizational Economics*. Princeton University Press: Princeton, NJ, US.

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

Long, R.J. and Fang, T. 2012. Do employees profit from profit sharing? Evidence from Canadian panel data. *ILR Review* 65(4), 899–927.

Longhi, S., Nijkamp, P. and Poot, J. 2006. Spatial heterogeneity and the wage curve revisited. *Journal of Regional Science* 46(4), 707–731.

Lucifora, C. 2015. Performance-related pay and labor productivity. *IZA World of Labor* No. 152.

Macpherson, D.A. and Hirsch, B.T. 1995. Wages and gender composition: Why do women's jobs pay less?. *Journal of Labor Economics* 13(3), 426–471.

Maczulskij, T. 2013. Public–private sector wage differentials and the business cycle. *Economic Systems* 37(2), 284–301.

Medoff, J.L. and Abraham, K.G. 1980. Experience, performance, and earnings. *The Quarterly Journal of Economics* 95(4), 703–736

Medoff, J.L. and Abraham, K.G. 1981. Are those paid more really more productive? The case of experience. *The Journal of Human Resources* 16(2), 186–216.

Moers, F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30(1): 67–80.

Moore, W.J., Newman, R.J. and Turnbull, G.K. 1998. Do academic salaries decline with seniority? *Journal of Labor Economics* 16(2), 352–366.

Mulligan, C.B. and Rubinstein, Y. 2008. Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics* 123(3): 1061–1110.

Murphy, K.J. 1992. Performance measurement and appraisal: motivating managers to identify and reward performance. In: W.J. Bruns (Ed.) *Performance measurement, evaluation, and incentives*, pp. 37–62. Boston, MA, US: Harvard Business School Press.

Neal, D.A. and Johnson, W.R. 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104(5), 869–895.

Nijkamp, P., and Poot, J. 2005. The last word on the wage curve?. *Journal of Economic Surveys* 19(3): 421–450.

Paarsch, H.J. and Shearer, B. 2000. Piece rates, fixed wages, and incentive effects: Statistical evidence from payroll records. *International Economic Review* 41(1), 59–92.

Pekkarinen, T. 2001. The wage curve: Evidence from the Finnish metal industry panel data. *Finnish Economic Papers* 14(1), 51–60.

Pekkarinen, T. and Riddell, C. 2008. Performance pay and earnings: Evidence from personnel records. *ILR Review* 61(3), 297–319.

Pekkarinen, T. and Uusitalo, R. 2012. Aging and productivity: Evidence from piece rates. *IZA Discussion Paper* No. 6909.

Pekkarinen, T. and Vartiainen, J. 2006. Gender differences in promotion on a job ladder: Evidence from Finnish metalworkers. *ILR Review* 59(2), 285–301.

Pema, E. and Mehay, S. 2010. The role of job assignment and human capital endowments in explaining gender differences in job performance and promotion. *Labour Economics* 17(6), 998–1009.

Pergamit, M.R., and Veum, J.R. 1999. What is a promotion?. *ILR Review* 52(4): 581–601.

Petersen, T., Snartland, V. and Meyersson Milgrom, E.M. 2007. Are female workers less productive than male workers? *Research in Social Stratification and Mobility* 25(1): 13–37.

Piekkola, H. 2005. Performance-related pay and firm performance in Finland. *International Journal of Manpower* 26(7/8), 619–635.

Phelps, E.S. 1972. The statistical theory of racism and sexism. *American Economic Review* 62(4), 659–661.

Polachek, S.W. 2014. Equal pay legislation and the gender wage gap. *IZA World of Labor* No. 2014: 16.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37(1), 7–63.

Prendergast, C. and Topel, R. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37(2–3): 355–65.

Sato, Y. 2000. Search theory and the wage curve. *Economics Letters* 66(1), 93–98.

Shapiro, C. and Stiglitz, J.E. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74(3), 433–444.

Shapiro, D.M. and Stelcner, M. 1987. The persistence of the male-female earnings gap in Canada, 1970-1980: The impact of equal pay laws and language policies. *Canadian Public Policy/Analyse de Politiques* 13(4), 462–476.

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

Solon, G., Barsky, R. and Parker, J.A. 1994. Measuring the cyclicality of real wages: How important is composition bias?. *The Quarterly Journal of Economics* 109(1), 1–25.

Statistics Finland. 2016. *Statistical Yearbook of Finland 2016*. Helsinki, Finland.

Topel, R. 1991. Specific capital, mobility, and wages: wages rise with job seniority. *Journal of Political Economy* 99(1), 145–176.

Toutkoushian, R.K. 1998. Sex matters less for younger faculty: evidence of disaggregate pay disparities from the 1988 and 1993 NCES surveys. *Economics of Education Review* 17(1), 55–71.

Uusitalo, R. and Vartiainen, J. 2009. Finland: Firm factors in wages and wage changes. In: Lazear, E.P. and Shaw, K.L. (ed.) *The Structure of Wages: An International Comparison*. University of Chicago Press: Chicago, IL, US.

Waldfogel, J. 1998. The family gap for young women in the United States and Britain: Can maternity leave make a difference?. *Journal of Labor Economics* 16(3), 505–545.

Ward, M. 2001. Gender and promotion in the academic profession. Scottish *Journal of Political Economy* 48(3), 283–302.

Weibel, A., Rost, K. and Osterloh, M. 2010. Pay for performance in the public sector—benefits and (hidden) costs. *Journal of Public Administration Research and Theory* 20(2), 387–412.

Weiss, A. 1995. Human capital vs. signalling explanations of wages. The *Journal of Economic Perspectives* 9(4), 133–154.

Yankow, J.J. 2006. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics* 60(2), 139–161.

CHAPTER 2: WORKER OUTPUT AND SENIORITY IN PERFORMANCE EVALUATIONS AND PROMOTIONS *

Abstract
This study employs unique longitudinal personnel data from a large university to examine the importance of worker output and job seniority in subjective performance evaluations and promotion decisions. A descriptive analysis of the data demonstrates that (1) employees higher up the job ladder received better subjective performance grades, on average, (2) within-job level performance grades tended to cluster into a few "norm" grades and (3) downgraded performance ratings and demotions along the job ladder were infrequent, suggesting that nominal wage cuts were rare. The econometric analysis shows that better-performing employees – with output measured both in absolute terms and relative to peers – were more likely to be assigned higher performance grades, indicating that subjective performance evaluations reflected actual differences in worker output. Furthermore, the results suggest that the probability of being promoted to higher job levels was positively associated with relative and absolute work output and that promotion probability increased with prior subjective performance grades even after accounting for recent work output. Additionally, the findings suggest that employees with more job seniority were evaluated as exhibiting higher performance than their equally productive but less-experienced peers. Finally, the results provide some evidence that both the probability of being upgraded to a higher performance grade and the probability of being promoted along the job ladder increased with years of job seniority, although at a diminishing rate.

Keywords: performance evaluation, promotions, job seniority, worker output
JEL Classification: J24, M51, M52

---

* This chapter is a joint work with Jaakko Pehkonen. The financial support from the Finnish Cultural Foundation (grant number 00120284) and the OP Group Research Foundation (grant number 201500181) is gratefully acknowledged.

1 Introduction

To generate incentives for employees to exert greater work effort, employers may adopt compensation schemes that link wage bonuses or promotions to higher-paying jobs to employee performance (e.g., Lazear and Rosen 1981; Lucifora 2015). While employers can sometimes tie pay and promotion decisions on observed measures of worker performance, such as units of output produced or sales records, the use of such measures as a basis for employee rewards may be impractical in working environments where all tasks are not be easily observable or measureable or where measures are inaccurate or vary for reasons unrelated to work effort (e.g., Holmström 1979; Holmström and Milgrom 1991; Baker et al. 1994). As an alternative or a complement to tying rewards to observable performance measures, employers may utilize subjective evaluations of employees' work performance to determine pay increases and promotions. While subjective performance evaluations potentially offer a more comprehensive view of job performance by accounting for work aspects that are not easily quantified and measured (Prendergast and Topel 1993), they also leave room for a supervisor's personal opinions, impressions and preferences, which may lead to biased evaluation outcomes (Bol 2008, 2011; Bol and Smith 2011). Empirical studies report several such biases, including supervisors' tendency to refrain from differentiation among employees by clustering performance ratings around certain "norm" ratings (e.g., Murphy 1992) or to show leniency in performance appraisals by rating subordinates higher than their actual performance warrants (Jawahar and Williams 1997; Moers 2005).

The limitations and noisiness of both observable and subjective performance measures make it difficult for employers to identify those employees who currently outperform their peers, and who should thus be rewarded with wage bonuses or promotions to higher paying jobs. When employers have limited or inaccurate information on worker performance, years of experience with the current employer may provide a useful proxy for worker productivity in reward decisions, given that workers with more job seniority may have accumulated productivity-enhancing skills by means of learning-by-doing and on-the-job training (e.g., Hutchens 1989; Topel 1991; Altonji and Williams 2005). Furthermore, years of job seniority potentially carry information on past work performance if employers offer longer labor contracts for better-performing employees. However, job seniority-related rewards may also arise from supervisory biases if supervisors favor senior workers for reasons unrelated to their work effort (for example due to closer interpersonal relationships) and hence reward them with better performance evaluations, pay increases or promotions even when their actual work performance does not warrant those rewards (e.g., Prendergast and Topel 1996).

In this paper, we employ longitudinal personnel data from a large university to provide new insights into how supervisors utilize information on job seniority and employees' work output – measured both in absolute terms and relative to peers – to undertake performance appraisals and make promotion decisions. The personnel data provide unique information on worker performance and job characteristics that allow us to overcome some important limitations of past studies. We first examine the relationship between performance grades and a set of worker-specific output measures that capture a wide range of academic duties (hereafter referred to as "objective performance measures") to assess whether subjective evaluations of performance provide accurate information regarding actual work performance differences. Empirical analyses of the relationship between subjective and objective measures of job performance are limited by the scarcity of suitable data sets. Previous empirical evidence on this relationship, particularly from the research literature in personnel psychology, have relied mainly on simple correlation analyses (e.g., Bommer et al. 1995; Heneman 1986) and experimental designs in which participants evaluated the job performance of hypothetical employees under experimenter-controlled conditions (e.g., Bol and Smith 2011). We use an econometric analysis of worker-specific performance data obtained from a real-life evaluation setting in which performance ratings were used in actual merit pay decisions to assess the relationship between subjective and objective measures of work performance.

Next, we assess whether the probability of being promoted along the job ladder depended on employees' work performance and job seniority. In the absence of data on worker output, past studies on promotion determinants have often used subjective performance evaluations (e.g., Flabbi and Ichino 2001; DeVaro 2006; DeVaro and Waldman 2012) or performance-related pay components (e.g., Cassidy et al. 2016; DeVaro and Kauhanen 2016) to control for productivity differences among workers[1]. The personnel data used in this study allow us to analyze whether promotion decisions were based on objective and subjective measures of work performance when accounting for differences in worker and job characteristics. Furthermore, past studies attempting to identify the predictors of promotions have been complicated by uncertainty regarding the precise hierarchy of jobs, making it difficult to distinguish promotions from other within-firm job changes such as formal upgrades of a current position that do not involve changes in job duties (e.g., Pergamit and Veum 1999). The university analyzed in this study uses an accurate job hierarchy based on a complexity ladder of job duties in which higher levels are associated with greater complexity, responsibility and autonomy, allowing us to examine promotions along a well-defined job ladder.

The analysis yields empirical evidence to answer the following research questions: Do subjective performance evaluations of employees reflect actual differences in worker output? Are senior workers more likely to be rewarded with higher performance grades and promotions than their less experienced

1 These studies generally find that better-performing employees have a higher probability of being promoted than their less well-performing colleagues with equal qualifications. A recent study by Frederiksen et al. (2017) combines personnel data from several previously published studies to provide consistent evidence that subjective evaluations of work performance are positively related to promotions.

colleagues with equal qualifications and work output? Are better-performing workers – measured using both objective and subjective measures – more likely to be promoted?

The remainder of the paper is organized as follows. Section 2 describes the institutional background of the study, consisting of a brief description of the Finnish university system and a characterization of the compensation system used in Finnish academia. Section 3 describes the personnel data and summarizes the empirical results. Finally, Section 4 concludes with a summary of the key findings.

2 Finnish universities, academic salaries and employee evaluation

2.1 Universities and the salary system

There are 14 universities in Finland: ten are multidisciplinary, and four specialize in arts, technology or business[2]. Universities enjoy autonomy in research and teaching and can independently organize their internal administration and organizational structure within the boundaries established by the relevant legislation (Universities Act). The universities are primarily funded by the central government, and their core funding comes from the state budget. The Ministry of Education and Culture determines the key parameters of the funding model for the universities. In essence, the model allocates funding based on university-specific teaching and research output measures, which determine approximately three-quarters of each university's state budget-based core funding. The key output measures of the funding model include (1) the number of completed bachelor's, master's and doctor's degrees, (2) the number of enrolled students who achieve a threshold value of 55 credits each academic year, (3) a quality-weighted count of scientific publications and (4) the amount of research funding received from external sources. In addition to output measures, one-quarter of the public core funding is based on university-specific factors (e.g., the composition of academic disciplines) and strategic considerations that aim, for example, to promote structural reforms that strengthen the universities' areas of expertise. In 2015, the state budget for universities was approximately 1.9 billion euros. In addition to direct state funding, the universities obtained approximately 0.63 billion euros from external funding sources, including public and private foundations, governmental funding bodies and private-sector firms. At

---

2 Based on the "*QS World University Rankings 2016–2017*", ten Finnish universities ranked among the top 550 universities in the world, with the University of Helsinki ranked highest (91st). The university analyzed in this study, the University of Jyvaskyla, ranked 338th.

the end of 2015, the universities employed 14,917 full-time faculty members and 9,873 administrative and other full-time staff members.[3]

The salary system in academia – which is uniform across all Finnish universities – relates employee remuneration to job complexity and personal work performance[4]. Based on the complexity of job duties and responsibilities, each faculty member is assigned to one of the eleven complexity levels (or "rungs" of the complexity ladder). Faculty members hold numerous academic occupational titles that can be roughly categorized according to their job complexity[5]: early-career researchers working on their doctoral dissertation (e.g., PhD students and teaching assistants) work at complexity levels 1–4, with the exact level depending on progress in graduate studies; researchers who have recently completed their doctoral studies work at complexity level 5 and those who have more experience and are established postdoctoral researchers work at levels 6–7; full professors work at complexity levels 8–11.[6] In addition to these research-oriented occupations, the faculty includes two teaching-oriented occupations: university instructors who work at complexity levels 4–5 and university lecturers who work at complexity levels 5–7. The key point highlighted here is that the eleven-rung complexity ladder effectively determines the job hierarchy at the universities but not the occupational titles (academic ranks). For example, although full professor is the highest academic rank, each full professor is also assigned to one of the four within-rank complexity levels (8–11) based on the nature of his or her job tasks and responsibilities. In 2015, 39% of the full-time faculty members of the research universities were working at complexity levels 1–4, 46% were at levels 5–7 and 15% were at levels 8–11 (AFIEE 2015).

The job complexity ladder determines the salary scale for faculty, with base monthly salaries ranging from 1,795 euros at the lowest complexity level to 6,831 euros at the highest complexity level in 2015. The base monthly salaries attached to each rung of the complexity ladder are established by a collective bargaining agreement. Collective bargaining over salaries and working conditions is conducted annually or biannually between the representatives of university employers' and employees' associations[7]. In addition to complexity level-specific base salaries, the salary system uses a nine-grade performance rating to determine salary variation within the complexity levels. The lowest performance grade (1) indicates that an employee earns a monthly salary that is equal to the base salary of the given complexity level. Performance grades from 2 to 9 increase the complexity level-specific base salary by approximately 4, 10, 16, 22,

---

3 Data sources: Ministry of Education and Culture, the Finnish National Board of Education and the Association of Finnish Independent Education Employers.

4 See "*General collective agreement for universities*" for a detailed description of the salary system (downloadable at http://www.sivistystyonantajat.fi/tiedostopankki/158).

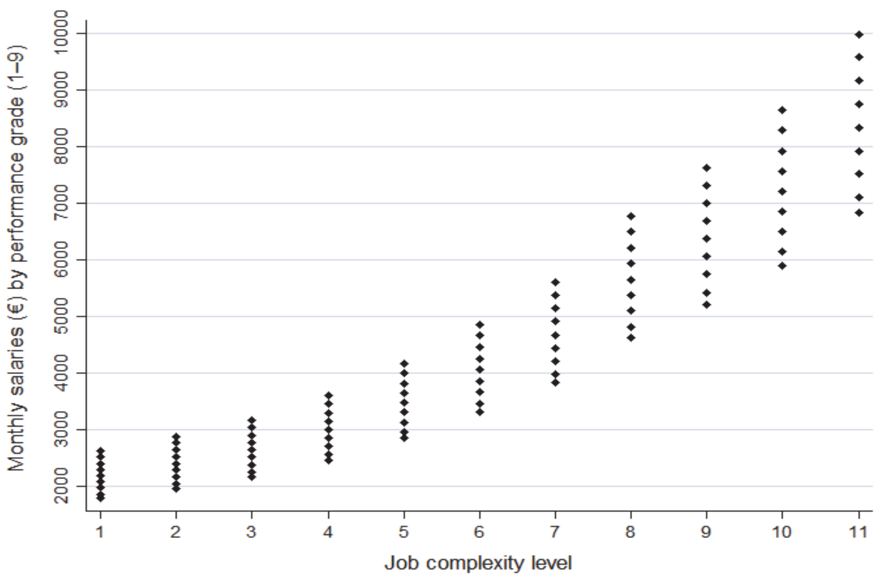
5 For example, in the university analyzed in this study, the faculty members had 23 distinct occupational titles in the period 2006–2012.

6 The research faculty occupations at complexity levels 5–7 correspond to the assistant and associate professor ranks of academia in North America and many European countries.

7 Over the 2006 through 2015 period, the collectively bargained base salaries increased, on average, by approximately 20%.

28, 34, 40, and 46 percent, respectively. Figure 1 reports monthly salaries by complexity level and performance grade in 2015. The figure reveals the overlapping nature of the salary system, as similar salary levels can be achieved with different combinations of complexity level and performance grade. Finally, faculty members can earn salary bonuses for some additional job tasks, most notably for administrative duties. In 2015, complexity level-specific base salary, performance grade-specific merit pay and other salary bonuses accounted for 81%, 17% and 2%, respectively, of an average employee's total monthly salary (AFIEE 2015).

FIGURE 1 Monthly salaries by complexity level and performance grade in 2015



2.2 Evaluation of job complexity and work performance

The collective agreement for the universities stipulates the general guidelines for the process used to evaluate employees' job complexity and work performance. However, the details of how the evaluations are conducted and are scheduled in practice may vary across universities. We will now describe in more detail the key aspects of the evaluation process, with an emphasis on the evaluation procedures of the case university examined in this study.

At the time of recruitment, each employee is assigned an immediate supervisor (typically the chair or vice-chair of a department) and classified into one of the eleven job complexity levels and into one of the nine performance grades based on their qualifications and prior merit. Following the initial classification, an employee's job complexity and work performance are assessed once every two years in a pre-scheduled assessment meeting between the immediate supervisor and the employee. The supervisor conducts the assessment meeting, makes an evaluation proposal and communicates the proposal to the employee. These job complexity and performance evaluations must be agreed upon and signed by the employee and the supervisor[8]. The evaluation outcome must then be approved by the dean of the faculty and finally by the central administration of the university, which assesses the consistency of (performance) evaluations across employees in the same discipline, occupation and job complexity level. The employee is entitled to request a re-assessment between the assessment meetings in the event of significant changes in his or her work performance, job duties or other working conditions. Similarly, the supervisor may request a re-assessment if a subordinate's job requirements and/or work performance have changed to such an extent that a new evaluation is considered necessary.

Evaluation of the employee's job complexity level is based on a job description that describes (1) the nature and the main responsibilities of the job, (2) the required interaction skills, and (3) the knowledge and skills needed for the job. A representative employee at the lowest complexity levels holds a master's degree, conducts postgraduate research and has small-scale teaching and administrative responsibilities, whereas a representative employee at the upper end of the complexity ladder holds a doctoral degree (with eligibility to serve as a full professor), has diversified teaching responsibilities with more advanced courses, supervises master's and doctoral theses, manages research projects and academic co-operation networks and has demanding administrative duties. In sum, a promotion along the job complexity ladder involves the diversification of job tasks and increased job qualification requirements, job responsibilities and job duty complexity.

The assessment of worker performance is based on an evaluation of individual success within the assigned complexity level, with success measured by achievements in three different activities: (1) teaching, (2) research and (3) societal engagement and contributions to the university community. Based on the instructions, employees' success in each of these activities is rated on a nine-grade scale, ranging from very low to excellent. Overall performance grade is the sum of the grades for the three different activities, weighted by the share of working time devoted to each activity.

Since the compensation is not explicitly linked to objective performance measures or years of university service, worker output and job seniority can influence salaries only indirectly by affecting supervisors' subjective judgments regarding whether employees should be awarded with higher performance grades or be promoted to higher rungs of the job complexity ladder. Furthermore, in addition to a direct wage effect of subjective performance evaluations,

---

8 The employee's option to appeal the evaluation outcome may potentially reduce subjective bias in assessment of work performance: to prevent appeals from unsatisfied employees, supervisors may put greater effort into evaluations to provide more accurate assessments (Prendergast and Topel 1993).

an indirect wage effect may arise if supervisors use information on employees' prior performance grades to guide their promotion decisions. Consequently, the pay scheme leaves significant scope for supervisory discretion in making compensation decisions. However, performance evaluations and promotions are guided by detailed administrative instructions, and supervisors conducting evaluations can easily access detailed information regarding their employees' recent work output (i.e., information-gathering costs are low), both of which may reduce subjective biases in compensation decisions (e.g., Bol 2011; Golman and Bhatia 2012).

Due to the formal salary system that links worker compensation to job complexity and personal performance, there are nine possible salary outcomes that can result from the assessment of these employee attributes. First, the salary level can remain unchanged if neither performance grade nor job complexity level is revised. Second, an employee's salary level can decrease in three cases: if complexity level and performance grade are lowered simultaneously or if only one of them is lowered. Conversely, an employee can receive a salary increase if either job complexity level or performance grade or both are increased. Third, the salary outcome is uncertain when the complexity level decreases and the performance grade increases, or vice versa.

3 Analysis of performance evaluations and promotions

3.1 The personnel data and empirical approach

The data analyzed in this study were obtained from the personnel records of a large Finnish multidisciplinary university for the years 2006–2012.[9] The personnel data were combined with worker-specific performance data drawn from an external database containing a variety of objective output measures. The original longitudinal data include information on all the teaching and research faculty of the university, and we restrict our analyses to only full-time employees in more established stages of their academic careers. Consequently, we exclude part-time workers and employees working at the lowest rungs of the job complexity ladder (1–4) from the final sample[10]. The final sample consists of 4,873 observations for 1,314 employees.

We first provide a descriptive analysis of the personnel data to identify some general features and patterns of performance evaluations and promotions. The descriptive analysis is followed by a more detailed scrutiny of the person-

---

9 A comparison of student and faculty characteristics suggests that the university analyzed in this study is representative of the multidisciplinary universities in Finland (see Table A1 in the Appendix).

10 Employees on the lowest rungs of the complexity ladder are mainly doctoral (PhD) students, whose compensation and career advancement is essentially based on a review of their progress in graduate studies.

nel data, and we use these data to estimate a set of regression models that relate subjective performance grades to worker characteristics and output measures. Next, we estimate linear probability models to examine the determinants of upward movement on a nine-grade performance rating scale and of promotions along the eleven-rung job complexity ladder. In the absence of adequate information on the reasons for employee attrition during the period under study, this analysis of career advancement decisions focuses on continuing employees (stayers).

The primary explanatory variables of interest in the estimated regression models are job seniority and a set of output measures extracted from the worker-specific performance data. Job seniority refers to length of service in the university, measured as years since entering the university[11]. The output measures reflect employees' work performance on a variety of academic duties. First, these measures include the yearly numbers of peer-reviewed articles (divided into international and domestic articles) and other publications (e.g., discussion papers, conference papers and book chapters). The number of peer-reviewed articles is likely to be the key determinant of compensation decisions because the state funding the university receives is partially linked to the number of these articles, offering supervisors strong incentive to emphasize this measure in personnel decisions. Furthermore, prior studies offer theoretical arguments suggesting that incentives for prioritizing research output may have strengthened in academia and present empirical evidence indicating that universities have become more inclined to base compensation and other personnel decisions essentially on research merits without considering other achievements (Laband and Tollison 2003; Remler and Pema 2009). If peer-reviewed articles are in fact the key determinant of compensation decisions and employees recognize this focus on research output, they may devote more of their working time to writing manuscripts for peer-reviewed journals at the expense of other job duties (see Holmström and Milgrom 1991; Baker 1992)[12].

In addition to the research output measures, the performance data contain information on a range of other academic output measures including yearly numbers of 1) bachelor's and master's theses and PhD dissertations supervised, 2) conference presentations and 3) additional academic activities and merits including honors, reviewer duties, editorships and fellowships, among others. Based on a preliminary analysis, we summed all these other output measures to produce an "other activities" variable. In contrast to the research output variables, which have fairly low mean values and variances, the values of the "other

---

11 In the original data set, some worker-year observations lack information on the years of university service or include information only on the length of time since the most recent labor contract was negotiated. However, the longitudinal nature of the data enabled us to impute many of the missing values and to correct for measurement error in the original job seniority variable, providing us with a corrected job seniority variable. In our analysis, we always employ this corrected job seniority variable.

12 Brickley and Zimmerman (2001), using data from a single business school, provide supporting evidence for this hypothesis by showing that faculty members responded to changes in incentives by devoting more of their work effort to job duties that became relatively more important determinants of compensation.

activities" variable are often substantially greater than zero and differ considerably among employees.

To summarize, we employ three variables to account for worker differences in job performance: 1) a count of peer-reviewed articles, 2) a count of other publications and 3) a count of "other activities". To focus on employees whose compensation is most likely to depend on these measures, the estimated regression models are based on a sample that excludes employees working in teaching-oriented occupations: university instructors and lecturers[13]. Furthermore, because the evaluation of worker performance may depend on workers' relative performance rather than their absolute performance – as the ability to distinguish how well an employee is performing depends essentially on the performance of the employee's peers – we also employ relative output measures in the analysis. The relative output measures are obtained by dividing the output measures for worker *i* during a given period by the average output measures of employees working in the same department and at the same complexity level as worker *i* in that period (hereafter, we use 'peers' to refer to employees who worked in the same department and at the same job complexity level in the same year). The relative output measures also account for discipline-related differences in worker output, especially in research output.

In addition to the job seniority and academic output measures, the estimated regression models include a set of control variables for worker characteristics. Employee age is included as a proxy for overall work experience. Furthermore, to account for the effects of gender, education and academic discipline, the models include dummy variables for female, highest degree completed (master's degree or lower, licentiate's degree, doctoral degree) and department[14].

Table 1 summarizes the data used in the analysis for all employees and for the most common job complexity levels. The table reveals several differences in worker characteristics and output across the complexity levels. First, employees higher up the complexity ladder were, on average, older and had more seniority than their colleagues on the lower rungs of the ladder. Furthermore, they were more often males and more likely to hold a doctoral degree. Finally, research output and the number of other activities typically increased with complexity level.

---

13 The data lack information on teaching loads and merits. However, there are two reasons that teaching performance might carry less weight in performance evaluations and promotions in the university analyzed in this study. First, teaching loads are typically uniform for employees in the same department within the same occupation (especially in occupations other than university instructor and lecturer). Second, assessing teaching skill is difficult because student evaluations of instructors are not collected.

14 In the final sample, 8.3% of the employee-year observations were missing data for the highest degree completed. We imputed the missing values with the most common degree of the employees working in the same occupation. However, the results of the estimated regression models remained essentially unchanged after excluding the employees with missing education data.

TABLE 1 Descriptive statistics (2006–2012)

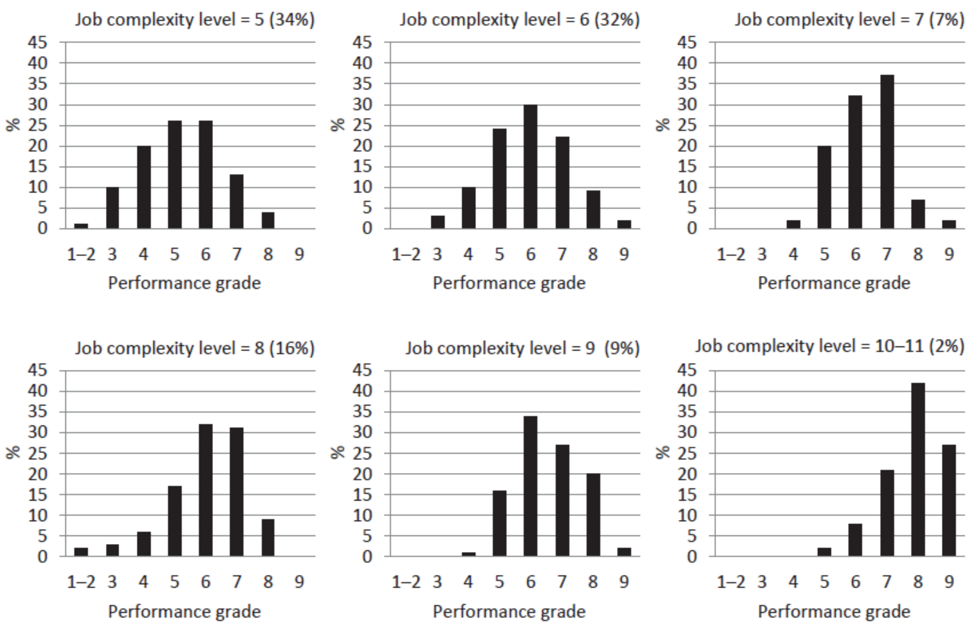
	All employees	Job complexity level 5	6	7	8	9
Job characteristics						
Job complexity level (5–11)	6.4					
	(1.5)					
Performance grade (1–9)	5.9	5.2	5.9	6.3	6.1	6.6
	(1.4)	(1.4)	(1.3)	(1.0)	(1.3)	(1.1)
Monthly earnings (euros)	4370	3272	3982	4671	5569	6530
	(1284)	(312)	(391)	(395)	(537)	(637)
Performance measures						
Refereed international articles	1.8	0.9	1.3	3.3	2.4	3.3
	(3.3)	(2.1)	(2.5)	(4.6)	(3.4)	(3.9)
Refereed national articles	0.3	0.2	0.3	0.5	0.6	0.5
	(1.0)	(0.7)	(0.8)	(1.3)	(1.3)	(1.0)
Other publications	1.2	0.6	1.0	1.7	1.9	2.3
	(2.3)	(1.5)	(1.9)	(2.9)	(2.8)	(3.0)
Other activities	9.9	3.3	8.5	14.2	17.0	21.5
	(12.5)	(6.5)	(10.4)	(12.3)	(13.5)	(16.4)
Worker characteristics						
Age (years)	46.8	41.7	46.6	50.0	51.3	54.0
	(9.9)	(9.6)	(9.1)	(7.5)	(8.6)	(6.8)
Job tenure (years in university)	11.1	7.7	11.5	13.5	12.8	15.6
	(9.5)	(7.5)	(9.0)	(10.6)	(10.4)	(9.8)
Gender (% males)	42.1	49.9	44.8	31.7	32.5	37.7
Doctoral (PhD) degree (%)	81.1	58.9	87.9	97.2	96.9	96.9
Career advancement						
Annual wage growth (%)						
All employees	3.7					
Promoted employees	15.2					
Non-promoted employees	3.1					
Promotion rate (yearly %)	5.0					
Performance grade increments (yearly %)	10.2					
Observations	4873	1673	1533	325	775	451

Notes: Table reports the average values of variables (standard deviations in parentheses). Promotion = increase of job complexity level (rank). Reported promotion rates are averages of yearly promotion rates. Promotion and earnings growth rates are based on employees who stayed at the university for (at least) two consecutive years.

3.2 Descriptive analysis of performance evaluations and promotions

Figure 2 reports the distributions of performance grades by job complexity level. Three conclusions can be drawn from the figure[15]. First, the performance grade distribution shifts to the right along the job complexity ladder, implying that employees in more complex jobs also had higher performance grades. The mean performance grade increases with each complexity level, from 5.2 (at complexity level 5) to 7.8 (at complexity levels 10–11). Tests for equality of means reveal that these mean differences are statistically significant at the 1% level. Furthermore, the figure illustrates that the lowest performance grades, 1 and 2, were very rarely used, which may be an indication of supervisor leniency. Finally, consistent with previous empirical findings (e.g., Flabbi and Ichino 2001; Dohmen 2004), the within-complexity level performance grades tend to cluster into a few "norm" grades, and the proportion of employees with one of the two most common performance grades varies from 52% (at complexity level 5) to 69% (at complexity levels 10–11). Because performance grades are attached to performance-based pay components, this compression of grades reduces wage variation within the complexity levels.

FIGURE 2 Performance grade distributions by job complexity level



[15] The figure aggregates data from all departments; however, the findings for individual departments are typically very similar to those reported here.

Table 2 illustrates the year-to-year job dynamics in the university by reporting movement along the complexity ladder and changes in performance grades. Four main observations can be discerned from this table. First, the standard outcome was "no change": complexity level and performance grade were unchanged in 84.3% of the cases. This observation is consistent with the evaluation instructions, which stress that a 'substantial change' in job requirements (work performance) must occur before an employee's position on the complexity ladder (performance grade) can be changed. Conversely, the observed stability of the complexity levels and performance grades may be at least partially attributable to budgetary considerations: because assignments to higher complexity levels or to higher performance grades result in higher salaries, supervisors may be inclined to avoid changes to employee evaluations to limit pay increases. The second observation from the table is that promotions along the complexity ladder were typically accompanied by a decrease in performance grade. Third, demotions along the complexity ladder were rare, with less than 1% of the cases resulting in an employee's demotion to a lower complexity level. The demotions often coincided with an increase in performance grade and, consequently, with an increase of the performance-based pay component (in approximately 60% of the cases), possibly due to supervisors' inclination to partially compensate for the decrease in base salary upon demotion. Fourth, performance grades were rarely downgraded, particularly when employees were not simultaneously promoted.

Table 3 reports the year-to-year transitions between job complexity levels. These results regarding movement along the complexity ladder are consistent with previous studies. First, the probabilities of remaining at the same complexity level were substantially higher than those of being promoted or demoted. As the table shows, the probability of promotion to a higher complexity level varied from 2.3% (at complexity level 9) to 11.6% (at complexity level 7). Second, transitions to lower complexity levels were rare, with probabilities varying from zero (at the highest complexity levels of 10–11) to 2.2% (at complexity level 8). Third, promoted employees typically moved one level up the job complexity ladder: one-level promotions accounted for approximately 94.6% of all promotions.

Table 4 reports the year-to-year transitions between performance grades, conditional on an unchanged complexity level. These probabilities, along with the transition probabilities in Table 2, imply that upgrades of performance grades were substantially more common than promotions along the complexity ladder. The probability of being upgraded to a higher performance grade varied from 30.4% (for performance grade 3) to 4% (for performance grade 8). Furthermore, the transition probabilities indicate that downgrades of performance rates were infrequent and were nonexistent at the upper end of the complexity ladder.

The observed rarity of demotions and reductions of performance grades suggests that salaries were rarely adjusted downward, which aligns with a number of empirical studies reporting that nominal wage cuts are uncommon (see, e.g., Gibbs and Hendricks 2004). The finding that supervisors refrain from personnel decisions that would result in wage cuts may be partially explained by supervisors' desire to avoid unwanted organizational outcomes. Prior research on the reasons for the downward rigidity of nominal wages suggests that employers (supervisors) refrain from cutting wages because of the negative effect wage cuts might have on employees' morale and loyalty and on the productivity and cohesion of the organization, as a consequence (Bewley 1995, 1998)[16].

TABLE 2 Changes of job complexity level and performance grade

		Job complexity level		
		Promotion	No change	Demotion
Performance grade	Increased	0.4	9.3	0.6 **
	No change	1.2	84.3	0.3 *
	Decreased	3.4	0.4 *	0.1 *

Notes: The number of observations is 3321. Shares are calculated for employees who stayed at the university for (at least) two consecutive years. Base monthly earnings decreased in * 100% of the cases; ** approximately 79% of the cases.

TABLE 3 Transition matrix between complexity levels

	Employee's complexity level at *t+1*					
Employee's complexity level at *t*	5	6	7	8	9	10–11
5	93.1	6.9	0	0	0	0
6	1.3	94.2	3.4	1.1	0	0
7	0	0.9	87.5	11.6	0	0
8	0	0.3	1.9	93.2	4.6	0
9	0	0	0	0.9	96.8	2.3
10–11	0	0	0	0	0	100

Notes: The number of observations is 3321. Transition probabilities are for employees who stayed at the university for (at least) two consecutive years.

16 Recent experimental findings are, in fact, consistent with employers' beliefs regarding the likely effects of wage cuts: Kube et al. (2013) found that wage cuts had a severe and persistent negative impact on worker productivity.

TABLE 4 Transition matrix between performance grades (conditional on an unchanged complexity level)

Employee's performance grade at *t*	Employee's performance grade at *t+1*							
	2	3	4	5	6	7	8	9
2	80.0	20.0	0	0	0	0	0	0
3	0	69.6	17.4	11.6	0	1.4	0	0
4	0	0	79.2	16.3	3.9	0.6	0	0
5	0	0.2	0.5	81.3	17.2	0.8	0	0
6	0	0	0.2	1.0	88.9	9.4	0.5	0
7	0	0	0	0	0	94.7	5.3	0
8	0	0	0	0	0	0	96.0	4.0
9	0	0	0	0	0	0	0	100

Notes: The number of observations is 3124. Transition probabilities are for employees (1) who stayed at the university for (at least) two consecutive years and (2) whose job complexity level was unchanged between *t* and *t+1*.

3.3 Determinants of performance evaluations

The previous section described the general patterns in the assignment of employees to different performance grades. We will now analyze the relationship between performance grades and worker characteristics (Tables 5 and 6) and the role of worker characteristics in determining upgrades to higher performance grades (Table 7).

Table 5 presents the results from regressing a dummy variable – which indicated whether an employee had an above-average performance grade among employees who were working in the same department and holding jobs with equal complexity – on relative job seniority, on relative output measures and on a set of control variables. Two main conclusions can be drawn from these results. First, the significant positive coefficient on the job seniority variable indicates that employees with more seniority were more likely to achieve the highest performance grades than their equally productive but less experienced peers (Column 3). This finding is robust to the inclusion of department dummies and observable worker characteristics including age, gender and education level (Column 4). The coefficient estimate from specification (4), which includes a full set of control variables, suggests that an employee with twice as much seniority as his or her peers was approximately 12% more likely to have an above-average performance grade.

Second, the results support a positive link between subjective performance ratings and relative worker output: employees with higher relative output were more likely to have above-average performance grades than their less productive peers with equally complex jobs. For example, based on the coefficient estimate in the last column of the table, an employee who published twice as many peer-reviewed articles as his or her peers with similar background characteristics was 8.4% more likely to have an above-average performance grade.

The observed positive relationship may indicate that supervisors were able to identify better-performing employees and to correctly assign them to higher performance grades. Conversely, the positive performance grade-output relationship may reflect reverse causality: higher performance grades (and the accompanying higher pay levels) may have had a motivational effect on the employees, resulting in increased worker output. The results in Table 5 link subjective performance evaluations with contemporaneous work output; however, using one-year lagged relative output measures as explanatory variables provided very similar results, with the exception that the coefficient on relative number of other publications was not statistically significant at the 10% level.

To more carefully examine the roles of job seniority and worker output in performance evaluation, Table 6 presents the results of the ordered probit analysis on the determinants of performance grades. All the specifications control for gender and education level and include dummy variables for departments and job complexity levels. The results from the specifications in columns (1) and (2) demonstrate that more senior and more productive employees – with seniority and worker output measured either in absolute terms or relative to peers – were more likely to have higher performance grades. To account for the fact that current performance grades are a function of past work output, the specification in the last column conditions on count variables that sum the output measures of the two previous years. The coefficient estimates on the count variables indicate that performance grades were significantly positively related to past worker output. To conclude, Table 6 lends further support to our conclusion from Table 5, which suggested that subjective performance grades reflect real differences in worker output.

To summarize, the results of Tables 5 and 6 provide strong evidence that workers with more job seniority were, on average, evaluated as exhibiting higher work performance than their less experienced colleagues with equally complex job duties and similar academic output and background characteristics. Consequently, because the pay scheme of the university links salaries to performance grades, this finding implies that senior workers earned a seniority wage premium. This finding accords with the positively sloped wage-seniority profile widely documented in the empirical literature on the determinants of wage levels (e.g., Dohmen 2004; Altonji and Williams 2005)[17].

---

17 In the organization analyzed here, wage level is explicitly tied to subjective performance ratings. However, prior empirical studies show that higher performance evaluations are also associated with higher wages in the absence of such explicit contracting (e.g., Dohmen 2004).

TABLE 5 Determinants of the highest performance grades

Dependent dummy variable = 1 if worker *i*'s performance grade in period *t* exceeded the average performance grade of employees working in the same department and at the same complexity level in period *t* as worker *i*; = 0 otherwise.

	(1)	(2)	(3)	(4)
Relative job seniority	0.081		0.072	0.061
	(0.021)***		(0.013)***	(0.023)***
Rel. number of refereed articles		0.035	0.039	0.042
		(0.011)***	(0.009)***	(0.011)***
Rel. number of other publications		0.016	0.015	0.014
		(0.006)***	(0.006)***	(0.006)**
Rel. number of other activities		0.080	0.070	0.071
		(0.012)***	(0.010)***	(0.013)***
Control variables				
Age, gender, highest degree	No	No	No	Yes
Department	No	No	No	Yes
$R^2_{adjusted}$	0.01	0.04	0.05	0.05
Observations	3134	2651	2651	2651

Notes: Coefficient estimates from OLS regressions (linear probability model), cluster-robust standard errors in parentheses (clustered at the worker level). Estimations are based on a sample that excludes 1) lecturers and university instructors and 2) those employees who had no peers working in the same department and at the same complexity level in the same period. Relative variables are calculated as follows: (The value of the variable for worker *i* in period *t*) / (The average value of the variable in period *t* for employees working in the same department and at the same complexity level as worker *i*). All models include a constant term. Full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

TABLE 6 Ordered probit on performance grade

	Relative output (t)	Absolute output (t)	Absolute output ($t{-}2$ + $t{-}1$)
Relative job seniority	0.147***		
Job seniority		0.065***	0.075***
Job seniority^2		−0.002***	−0.002***
Worker output (relative)			
Refereed articles	0.076***		
Other publications	0.028**		
Other activities	0.180***		
Worker output (absolute)			
Refereed articles		0.032***	0.024***
Other publications		0.044***	0.043***
Other activities		0.013***	0.008***
Control variables			
Gender, highest degree	Yes	Yes	Yes
Department	Yes	Yes	Yes
Complexity level dummies	Yes	Yes	Yes
Sample	*Total sample*	*Total sample*	*Those who had worked for at least three consecutive years*
Observations	2651	3311	1391
Pseudo R^2	0.11	0.12	0.15
Log pseudolikelihood	−4096	−5097	−1965

Notes: Table reports the ordered probit coefficients. Estimations are based on a sample that excludes lecturers and university instructors. Relative variables are calculated as follows: (The value of the variable for worker i in period t) / (The average value of the variable in period t for employees working in the same department and at the same complexity level as worker i). Estimations that include relative variables exclude those employees who had no peers working in the same department and at the same complexity level in the same period. Standard errors (clustered at the worker level) are not reported here due to space considerations. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level. Full results are available upon request.

Table 7 presents results from estimating a set of linear probability models of year-to-year performance grade increments, with the dependent variable indicating whether or not an employee's performance grade increased from the previous year. To account for the fact that the performance grade changes of employees who moved up or down the complexity ladder were likely less related to job performance than to other considerations (as illustrated in Table 2, promotions were typically associated with a decrease in performance grade, and demotions were associated with an increase in performance grade), the estimated models are based on a sample of employees whose complexity levels remained unchanged from the previous year. All the specifications control for initial performance grade and a full set of worker characteristics (including age, gender, education level and department).

The results suggest that the probability of being promoted to a higher performance grade was not related to relative job seniority (column 1) but that it increased with years of university service, although at a diminishing rate (column 2). The specifications in columns (3) and (4) incorporate worker-specific output measures. The seniority estimate is changed very little by the inclusion of relative output measures (column 3). However, when absolute output measures are used to control for recent worker output (column 4), the job seniority coefficient decreases and is no longer significant, suggesting that the job seniority and absolute output measures are correlated. The coefficient estimates for the output measures indicate that the probability of upward movement on the performance grade scale increased with relative (column 3) and absolute (column 4) output of peer-reviewed articles and other academic activities (e.g., theses supervisions, reviewer duties and conference presentations)[18]. Interestingly, the results indicate that a higher output of other (non-peer reviewed) publications relative to peers decreased the probability of being upgraded to a higher performance grade.

18 To test the robustness of these results to alternative output measures that were measured over a longer period, we used a sample of employees who had worked in the university for at least three consecutive years to re-estimate the specification in column (4) with count measures that summed the values of the output measures of the two previous years. The results were qualitatively similar to those reported in column (4), except that the coefficient estimates on the output measures were reduced by roughly half. The results of these estimations are available upon request.

TABLE 7 Increase of performance grade (conditional on an unchanged complexity level)

Dependent dummy variable = 1 if performance grade increased between *t−1* and *t*; = 0 otherwise

	(1)	(2)	(3)	(4)
Relative job seniority$_{t-1}$	−0.008			
Job seniority$_{t-1}$		0.005**	0.006**	0.003
Job seniority$^2_{t-1}$		−0.0002**	−0.0002***	−0.0001*
Worker output (relative)				
Refereed articles$_{t-1}$			0.024***	
Other publications$_{t-1}$			−0.009**	
Other activities$_{t-1}$			0.018**	
Worker output (absolute)				
Refereed articles$_{t-1}$				0.027***
Refereed articles$^2_{t-1}$				−0.001***
Other publications$_{t-1}$				−0.009
Other publications$^2_{t-1}$				0.0007
Other activities$_{t-1}$				0.004***
Other activities$^2_{t-1}$				−0.00003**
Control variables				
Performance grade dummies (t−1)	Yes	Yes	Yes	Yes
Department dummies (t−1)	Yes	Yes	Yes	Yes
Age, gender, highest degree (t−1)	Yes	Yes	Yes	Yes
$R^2_{adjusted}$	0.03	0.04	0.05	0.07
Observations	1902	2014	1630	2014

Notes: Coefficient estimates from OLS regressions (linear probability model). Estimations are based on a sample that excludes lecturers and university instructors. Relative variables are calculated as follows: (The value of the variable for worker *i* in period *t*) / (The average value of the variable in period *t* for employees working in the same department and at the same complexity level as worker *i*). Estimations that include relative variables exclude those employees who had no peers working in the same department and at the same complexity level in the same period. All models include a constant term. Standard errors (clustered at the worker level) are not reported here due to space considerations. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level. Full results are available upon request.

3.4 Determinants of promotions

We will next examine the factors affecting the probability of being promoted along the complexity ladder. In the estimated linear probability models, promotion refers to an increase in an employee's job complexity level between consecutive years. We argue that the complexity ladder provides a relevant way to define the job hierarchy of the university (and an increase of complexity level represents a relevant way to define promotion) because movement up the complexity ladder is accompanied by greater complexity, responsibility and autonomy. In addition to controls for worker characteristics (age, gender, highest degree and department), all the estimated specifications control for initial complexity level to account for the limited number of complexity levels and the possibility that promotion probabilities differ across complexity levels.

The main results of the estimated promotion models are reported in Table 8. The coefficient estimates for the job seniority variables can be summarized as follows: (1) promotion probability was not related to an employee's job seniority relative to peers and (2) promotion probability increased with years of job seniority but at a diminishing rate; however, the positive seniority effect is statistically significant only in the specifications that do not include controls for prior performance grades, suggesting that years of job seniority and performance grades are correlated. The results for the output measures illustrate that the probability of being promoted to higher complexity levels was significantly positively related to the relative and absolute output of peer-reviewed articles and other activities. Conversely, and consistent with the results obtained for performance grade increments, the estimates consistently suggest that employees who had a higher output of other publications – measured both in absolute terms and relative to peers – were statistically significantly less likely to be promoted, implying that production of non-peer reviewed publications was penalized with reduced promotion opportunities. For example, the coefficient estimate in specification (6) suggests that an additional non-peer reviewed publication decreased promotion probability by approximately 1%.

The specifications in columns (3) and (6) are also conditioned on prior performance grades, whereas the specification in column (4) adds a dummy variable indicating whether an employee had an above-average performance grade among employees who worked in the same department and in equally complex jobs. The results illustrate that opportunity for promotion depended on both prior performance grades and recent work output. Hence, consistent with previous studies (e.g., Flabbi and Ichino 2001; Dohmen 2004), the results suggest that information on prior subjective performance evaluations was involved in determining promotions, as employees with better performance grades were more likely to be promoted to higher rungs on the job ladder. Our results demonstrate that this finding is robust to the inclusion of controls for worker output (measured either in relative or absolute terms), implying that prior subjective performance evaluations play a role in promotions even after accounting for differences in recent output, suggesting that a better performance evaluation history increases the probability of being promoted regardless of actual work output.

TABLE 8 Promotions along the job complexity ladder

Dependent variable = 1 if complexity level increased between $t-1$ and t; = 0 otherwise						
	(1)	(2)	(3)	(4)	(5)	(6)
Relative job seniority$_{t-1}$	0.006					
Job seniority$_{t-1}$		0.004**	0.002	0.003	0.004**	0.002
Job seniority$^2_{t-1}$		−0.0002***	−0.0001	−0.0001**	−0.0001***	−0.0001
<u>Worker output (relative)</u>						
Refereed articles$_{t-1}$	0.015***	0.014**	0.011*	0.012**		
Other publications$_{t-1}$	−0.007**	−0.007**	−0.008**	−0.008**		
Other activities$_{t-1}$	0.013*	0.013*	0.006	0.009		
<u>Worker output (absolute)</u>						
Refereed articles$_{t-1}$					0.014***	0.011***
Refereed articles$^2_{t-1}$					−0.0007***	−0.0005**
Other publications$_{t-1}$					−0.009***	−0.011***
Other publications$^2_{t-1}$					0.0003*	0.0004**
Other activities$_{t-1}$					0.003***	0.002*
Other activities$^2_{t-1}$					−0.00003**	−0.00002
<u>Performance grade</u>						
Top performer$_{t-1}$				0.051***		
Perf. grade$_{t-1}$ = 1 or 2			−0.082***			−0.001
Perf. grade$_{t-1}$ = 3			−0.084***			−0.090***
Perf. grade$_{t-1}$ = 4			−0.065***			−0.056***
Perf. grade$_{t-1}$ = 5			−0.032**			−0.031**
Perf. grade$_{t-1}$ = 6			(*ref*)			(*ref*)
Perf. grade$_{t-1}$ = 7			0.032*			0.023
Perf. grade$_{t-1}$ = 8			0.101***			0.126***
Perf. grade$_{t-1}$ = 9			0.186**			0.142*
Control variables (t−1)	Complexity level, department, age, gender, highest degree					
$R^2_{adjusted}$	0.03	0.03	0.06	0.04	0.04	0.06
Observations	1744	1744	1744	1744	2155	2155

Notes: Coefficient estimates from OLS regressions (linear probability model). Standard errors (clustered at the worker level) are not reported here due to space considerations. Estimations are based on a sample that excludes lecturers and university instructors. Relative variables are calculated as follows: (The value of the variable for worker *i* in period *t*) / (The average value of the variable in period *t* for employees working in the same department and at the same complexity level as worker *i*). Estimations that include relative variables exclude those employees who had no peers working in the same department and at the same complexity level in the same period. Dummy-variable *Top-performer* = 1 if worker *i*'s performance grade in period *t* exceeded the average performance grade of employees working in the same department and at the same complexity level in period *t* as worker *i*; = 0 otherwise. All models include a constant term. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level. Full results are available upon request.

4 Conclusions and discussion

This study utilizes a longitudinal data set drawn from the personnel records of a large multidisciplinary university to examine the importance of worker output and job seniority as predictors of employee performance evaluations and promotions. This unique data set contains detailed information on worker attributes not typically available in administrative data sets, which allows us to address some of the limitations of previous studies of the same topic. First, the data set contains information on both *subjective performance evaluations* and *objective performance measures*, allowing us to directly estimate the relationship between subjective evaluations of worker performance and actual work performance. Furthermore, the data set provides a detailed *complexity ladder of job duties* that enables us to overcome challenges related to the determination of job hierarchy, allowing us to analyze promotions along a well-defined job ladder.

A descriptive analysis of the personnel data leads to several interesting conclusions regarding employee performance evaluations and promotions. First, the mean performance grade increased along the job ladder, indicating that employees higher in the job hierarchy also tended to receive higher performance grades. Second, the analysis illustrates that within-job level performance grades tended to cluster into a few "norm" grades. This observation, together with the finding that the lowest and highest performance grades were rarely used, aligns with findings from previous studies (e.g., Flabbi and Ichino 2001; Dohmen 2004) that supervisors tend to compress employee performance ratings. This performance evaluation compression may stem in part from supervisors' efforts to control payroll costs or promote salary equality (e.g., Prendergast and Topel 1993). Third, the findings suggest that there was notable downward rigidity in the subjective performance evaluations, as performance grades were almost never downgraded. Supervisors may have refrained from lowering performance grades for a number of reasons, such as not wanting to communicate poor evaluations to employees, to prevent the possible demotivating effects of lower evaluations or to motivate employees who were currently performing poorly (e.g., Murphy 2008). Fourth, the descriptive analysis reveals that demotions along the job complexity ladder were rare.

The observed rarity of personnel decisions with negative salary consequences (i.e., reductions of performance grades and demotions) may at least partially result from the attrition of poor performers. While our study focuses only on the performance grade and job level transitions of continuing employees (i.e., employees who remained employed for at least two consecutive years), it is possible that employees with deteriorating performance were more likely to leave the university, thereby potentially explaining the observed absence of employee evaluation outcomes that would entail salary cuts. Unfortunately, in the absence of specific information on reasons for leaving the university, we cannot test this hypothesis in more detail. However, the infrequency of employee evaluations that negatively affected salaries for continuing employees suggests that nominal wage cuts were rare. The downward rigidity of nominal wages is a widely documented empirical phenomenon (see, e.g., Gibbs and Hendricks 2004). An interview study by Bewley (1995, 1998) sheds some light on the reasons for downward wage rigidity, suggesting that supervisors refrain from nominal wage cuts because they might impair employees' morale and loyalty to the organization.

The analysis of the determinants of performance evaluations reveals that when comparing employees with equally complex job duties and similar background characteristics, the employees who outperformed their peers (i.e., who had higher relative work output) were more likely to have above-average performance grades. Moreover, employees with higher (relative) work output tended to have higher performance grades than their equally qualified colleagues with lower output. These findings indicate that subjective performance grades reflected actual differences in worker output, suggesting that subjective performance evaluations may provide an applicable proxy variable for worker output that might be effectively used in personnel decisions (e.g., pay raises, promotions, dismissals) and empirical research when direct measures of output are unavailable. Furthermore, because performance grades were attached to performance-related pay components, these findings illustrate that merit pay allocation was in fact based on actual worker output.

Additionally, the findings indicate that better-performing employees – with output measured both in absolute terms and relative to peers – had a higher probability of being upgraded to higher performance grades and being promoted to more complex jobs than their peers with similar characteristics but lower output. Since upward movement on the performance grade scale and on the complexity ladder were accompanied by a salary increase, this finding suggests that supervisors were able to recognize and reward good work performance. Furthermore, the promotion analysis illustrates that employees with higher performance grades had a higher probability of promotion, suggesting that supervisors also used prior evaluations of employee performance to guide their promotion decisions.

The analysis of performance evaluations also reveals that employees with more job seniority (measured in years of university service) tended to have higher performance grades than their less experienced peers with equal qualifications and output. Furthermore, the findings provide some evidence to suggest that more senior employees had a higher probability of being upgraded to higher performance grades and being promoted, although the statistical significance of the job seniority estimates depended on the included control variables. There are at least three plausible explanations for the positive relationship between job seniority and performance evaluations and promotions. First, job seniority may be positively correlated with (nonobservable or nonmeasurable) performance measures that are not included in the estimated regression models, such as the accumulation of productivity-enhancing skills with job seniority. Second, the higher performance grades – and the accompanying higher pay levels – of senior workers may also be a consequence of their good work performance in the past, presuming that justified performance grade increments in earlier years were not adjusted downward in later years, even if actual work performance deteriorated (see, e.g., Bishop 1987). Third, supervisors may be more lenient towards senior employees due to their stronger acquaintanceship ties with those employees (see, e.g., Lefkowitz 2000; Bol 2011).

References

AFIEE. 2015. *Tilastojulkaisu 2015: Yliopistot* (in Finnish). The Association of Finnish Independent Education Employers: Helsinki, Finland.

Altonji, J.G., and Williams, N. 2005. Do wages rise with job seniority? A reassessment. *Industrial and Labor Relations Review* 58(3), 370–397.

Baker, G. 1992. Incentive contracts and performance measurement. *Journal of Political Economy* 100(3), 598–614.

Baker, G., Gibbons, R. and Murphy, K.J. 1994. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics* 109(4), 1125–1156.

Bewley, T.F. 1995. A depressed labor market as explained by participants. *American Economic Review* 85(2), 250–254.

Bewley, T.F. 1998. Why not cut pay?. *European Economic Review* 42(3–5), 459–490.

Bishop, J. 1987. The recognition and reward of employee performance. *Journal of Labor Economics* 5(4), S36–S56.

Bol, J.C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* 27(1), 1–24.

Bol, J.C. 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review* 86(5), 1549–1575.

Bol, J.C., and Smith, S.D. 2011. Spillover effects in subjective performance evaluation: bias and the asymmetric influence of controllability. *The Accounting Review* 86(4), 1213–1230.

Bommer, W.H., Johnson, J.L., Rich, G.A., Podsakoff, P.M. and MacKenzie, S.B. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology* 48(3), 587–605.

Brickley, J.A., and Zimmerman, J.L. 2001. Changing incentives in a multitask environment: Evidence from a top-tier business school. *Journal of Corporate Finance* 7(4), 367–396.

Cassidy, H., DeVaro, J. and Kauhanen, A. 2016. Promotion signaling, gender, and turnover: New theory and evidence. *Journal of Economic Behavior & Organization* 126(A), 140–166.

DeVaro, J. 2006. Internal promotion competitions in firms. *The RAND Journal of Economics* 37(3), 521–542.

DeVaro, J. and Kauhanen, A. 2016. An "opposing responses" test of classic versus market-based promotion tournaments. *Journal of Labor Economics* 34(3), 747–779.

DeVaro, J. and Waldman, M. 2012. The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics* 30(1), 91–147.

Dohmen, T.J. 2004. Performance, seniority, and wages: formal salary systems and individual earnings profiles. *Labour Economics* 11(6), 741–763.

Flabbi, L. and Ichino, A. 2001. Productivity, seniority and wages: New evidence from personnel data. *Labour Economics* 8(3), 359–387.

Frederiksen, A., Lange, F. and Kriechel, B. 2017. Subjective performance evaluations and employee careers. *Journal of Economic Behavior & Organization* 134, 408–429.

Gibbs, M. and Hendricks, W. 2004. Do formal salary systems really matter? *Industrial and Labor Relations Review* 58(1), 71–93.

Golman, R. and Bhatia, S. 2012. Performance evaluation inflation and compression. *Accounting, Organizations and Societyi* 37(8), 534–543.

Heneman, R.L. 1986. The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology* 39(4), 811–826.

Holmström, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10(1), 74–91.

Holmström, B. and Milgrom, P. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization* 7(special issue), 24–52.

Hutchens, R.M. 1989. Seniority, wages and productivity: A turbulent decade. *Journal of Economic Perspectives* 3(4), 49–64.

Jawahar, I.M., and Williams, C.R. 1997. Where all the children are above average: The performance appraisal purpose effect. *Personnel Psychology* 50(4), 905–925.

Kube, S., Maréchal, M.A. and Puppe, C. 2013. Do wage cuts damage work morale? Evidence from a natural field experiment. *Journal of the European Economic Association* 11(4), 853–870.

Laband, D.N. and Tollison, R.D. 2003. Dry holes in economic research. *Kyklos* 56(2), 161–174.

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

Lefkowitz, J. 2000. The role of interpersonal affective regard in supervisory performance ratings: A literature review and proposed causal model. *Journal of Occupational & Organizational Psychology* 73(1), 67–85.

Lucifora, C. 2015. Performance-related pay and labor productivity. *IZA World of Labor* 152.

Moers, F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30(1), 67–80.

Murphy, K.J. 1992. Performance measurement and appraisal: motivating managers to identify and reward performance. In: W. J. Bruns (Ed.) *Performance measurement, evaluation, and incentives*. pp. 37–62. Boston, MA, US: Harvard Business School Press.

Murphy, K.R. 2008. Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology* 1(2), 148–160.

Pergamit, M.R. and Veum, J.R. 1999. What is a promotion? *ILR Review* 52(4), 581–601.

Prendergast, C. and Topel, R. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37(2–3): 355–65.

Prendergast, C. and Topel, R.H. 1996. Favoritism in organizations. *Journal of Political Economy* 104(5), 958–978.

Remler, D.K., and Pema, E. 2009. Why do institutions of higher education reward research while selling education? *NBER working paper series,* No. 14974.

Topel, R. 1991. Specific capital, mobility, and wages: wages rise with job seniority. *Journal of Political Economy* 99(1), 145–176.

Appendix

TABLE A1 University characteristics in 2014

	Other multidisciplinary universities (average values)	Our case university
Students and degrees [a]		
Number of students (bachelor's and master's levels)	10333	11613
Number of graduates (master's degree)	1103	1452
Number of graduates (PhD)	129	158
Faculty members		
Employees [a]	1254	1409
Share of female employees (%)	47	49
Earnings (euros)	3720 [b]	3703 [c]
Complexity level (min = 1, max = 11)	4.7 [b]	4.9 [c]
Journal articles/faculty members	0.95	0.89
Full professors		
Share of faculty members (%)	17	16
Earnings (euros)	6837	6831
Complexity level (min = 1, max = 11)	8.7	8.5

Notes: [a] Student and employee figures exclude the University of Helsinki (which had 28,185 students and 3941 faculty members); [b] Average values of earnings and complexity levels are calculated for faculty members of all Finnish universities (including the University of Jyvaskyla) in 2012; [c] Authors' own calculation from the personnel data (these figures are for the year 2012). Data sources: Vipunen-database (The Finnish National Board of Education), The Association of Finnish Independent Education Employers, The Finnish Union of University Professors.

CHAPTER 3: PROMOTIONS AND EARNINGS – GENDER OR MERIT? EVIDENCE FROM LONGITUDINAL PERSONNEL DATA *

Abstract
This study examines the determinants of promotions, performance evaluations and earnings using unique longitudinal data from the personnel records of a large university. The study focuses on the role of gender in remuneration using, first, information on the complexity ratings of job tasks to define promotions on job ladders and, second, information on objective individual productivity. The study finds that individual research productivity was an important determinant of promotions and earnings. The results indicate that gender has no effect on the probability of being promoted, conditional on productivity, nor does it play a role in the performance evaluation of employees. Furthermore, the results suggest that contemporaneous productivity measures provide a usable proxy for the past productivity of a worker.

Keywords: promotions, gender pay gap, individual productivity, performance evaluation, job complexity

JEL Classification: J16, J24, J31, M51, M52

---

* This chapter is a joint work with Jaakko Pehkonen. This chapter will be published in a forthcoming issue of the *Journal of Labor Research* (an online version of the article is available at https://doi.org/10.1007/s12122-017-9254-7). The authors thank Ari Hyytinen, Petri Böckerman and participants at the EALE Conference (Turin), the seminar of Labour Institute for Economic Research (Helsinki) and the Annual Meeting of the Finnish Economic Association (Mariehamn) for helpful comments. I acknowledge the financial support of the Yrjö Jahnsson Foundation (grant numbers 6217).

1 Introduction

An extensive empirical literature concludes that women earn less than otherwise identical men (e.g., Blau and Kahn 2000; Arulampalam et al. 2007). The literature indicates that barriers to promotion can explain a substantial part of the observed wage gap, as women are less likely to be promoted than are equally qualified men (e.g., Pergamit and Veum 1999). A significant shortcoming of existing literature is the inability to control for actual individual performance differences: observed gender differences in average earnings and promotion rates do not necessarily arise from discriminatory behavior but may simply reflect gender differences in worker output. Consequently, estimates of gender gaps in earnings, promotions, or both, may be biased by the omission of performance variables or by the use of biased proxy variables for performance, most notably, subjective performance evaluation ratings. Imperfect information about the hierarchy of jobs presents another difficulty for the analysis of promotion outcomes, as an ambiguous ranking of job titles within an organizational hierarchy complicates the identification of promoted workers. This problem is particularly evident in multi-organizational studies, as the range of job titles and their hierarchy can widely vary across organizations.

We present new evidence on the relative contributions of worker output and gender to promotions and earnings using longitudinal personnel data from a large Finnish university. The data set contains unique information on worker-specific productivity, employee performance evaluations and detailed job task complexity ratings, allowing us to analyze the role of gender in earnings determination, performance evaluations and promotion decisions within well-defined job ladders while accounting for differences in actual individual productivity.

2 Motivation and related literature

To assess whether gender plays a role in remuneration and promotion decisions, it is essential to compare the (average) pay and promotion rates of identical men and women who are performing equally well. Observed pay and promotion differences between men and women of similar merit and qualifications are conventionally interpreted as evidence of discriminatory behavior by employers, but they can also be ascribed to other factors, such as employee differences in negotiation skills and willingness to ask for pay increases and promotions (e.g., Booth 2009).

In the absence of data on individual output, studies on pay differences and promotion decisions conventionally use human capital–related proxy variables (such as tenure and education level) to control for potential productivity differences among workers. Additionally, some studies have used supervisors' performance evaluation scores of employees to proxy for actual productivity (e.g., Bartel 1995; Flabbi and Ichino 2001; Pekkarinen and Vartiainen 2006; Pema and Mehay 2010). The problem with this approach is that performance evaluations may be biased measures of actual productivity (Waldman and Avolio 1986; Prendergast and Topel 1993), most notably because supervisors tend to give more lenient and compressed evaluation ratings when they know that the ratings are used for administrative purposes (Jawahar and Williams 1997; Moers 2005). Moreover, the use of subjective performance evaluations to account for individual productivity differences is particularly problematic in the analysis of gender biases in earnings and promotions, as gender may be a significant determinant of performance evaluation scores (Bartol 1999; Castilla 2012).

In addition to the lack of worker productivity data, promotion studies are further complicated by the problem of defining the hierarchy of jobs: due to the wide variety of job titles, it can be difficult to identify which job changes within organizations should be regarded as promotions. To define promotions on job ladders, studies have typically deduced the job hierarchy from combined information on job titles, job descriptions and transitions between job titles (e.g., Baker et al. 1994; Dohmen et al. 2004). Alternatively, some studies have determined promotions using questionnaire information on self-reported job changes at the same employer (e.g., Francesconi 2001; Booth et al. 2003). The latter approach is potentially problematic because, as noted by Pergamit and Veum (1999), promotions reported by employees are not always actual promotions but mere formal upgrades of the current position that do not involve changes in job duties.

One particular labor market of highly skilled workers, namely, the academic labor market, provides an ideal setting for an analysis of career outcomes for two reasons. First, academia has a well-defined hierarchy of jobs, thereby facilitating the identification of promoted workers. Second, data on academic employees frequently include detailed individual performance measures, such as research productivity and teaching merit (e.g., Toutkoushian 1998, 1999; Monks and Robinson 2000). Previous empirical evidence suggests that academia is not an exception in regard to gendered remuneration: results from various countries – including the US (Toutkoushian 1998), the UK (Blackaby et al. 2005), Canada (Warman et al. 2010) and Japan (Takahashi and Takahashi 2011) – indicate that female academics earn less than male colleagues of comparable merit and productivity. Furthermore, gender pay inequality is evidently increased by gender-biased promotion procedures, as men in academia are more likely to be promoted than are women, even when conditioning on differences in individual qualifications and academic productivity (Ward 2001; Ginther and Hayes 2003). However, previous findings also illustrate that gender gaps in career outcomes are partly attributable to productivity differences, as the results show that the observed gender pay gap decreases when differences in academic achievements are considered (e.g., Barbezat 1991; Ransom and Megdal 1993).

The first contribution of our analysis is the use of detailed information on the complexity of job tasks to determine the hierarchy of jobs, allowing us to assess the roles of gender and productivity in promotions along well-defined job ladders. In contrast to some closely related promotion studies, including those by Pekkarinen and Vartiainen (2006), Van Herpen et al. (2006) and Kunze and Miller (2014), we use information on actual output rather than subjective performance evaluations to control for individual productivity differences. Furthermore, we contribute to the literature on employee performance appraisal by testing whether the gender gap in performance evaluations (Bartol 1999; Castilla 2012) is sensitive to the inclusion of variables measuring worker productivity. Finally, our results provide additional evidence of gender pay differences in formalized pay systems that partly tie compensation to worker performance. Such pay systems may reduce gender inequality in compensation for two reasons. First, the explicit guidelines of formalized wage systems may restrict supervisor discretion in pay and promotion decisions, leaving less room for gender discrimination (e.g., Elvira and Graham 2002). Second, performance-related compensation ought to limit the pay differences between male and female workers with similar outputs. However, the (indirect) empirical research on whether this is in fact the case is inconclusive: some findings suggest that the gender pay gap is smaller when workers are paid on the basis of output rather than on the time they spent working (Jirjahn and Stephan 2004; Petersen et al. 2007), while others indicate that gap is more pronounced in pay-for-performance wage systems (De la Rica et al. 2010; Kangasniemi and Kauhanen 2013).

Our analysis employs a longitudinal data set drawn from the personnel records of a single university (the University of Jyvaskyla). In recent decades, a growing body of empirical literature has utilized personnel data from single firms and universities to analyze the determinants of different career outcomes (e.g., Baker et al. 1994; Flabbi and Ichino 2001; Ransom and Oaxaca 2005; Haeck and Verboven 2012; Kelchtermans and Veugelers 2013; Dohmen et al. 2014). Although results based on a single organization should be interpreted with some caution, there are several advantages of using such data to study earnings and promotions decisions. First, the data from personnel records are typically highly accurate and contain detailed information not available in customary survey and administrative data sets, including worker-specific productivity measures and a well-defined hierarchy of job titles. Second, personnel data allow us to analyze earnings and promotion decisions within an internal labor market with homogenous personnel policies and uniform criteria for remuneration and career advancement. Third, as opposed to a multi-organizational study, we can ignore the effects of unobserved organization heterogeneity with respect to earnings (e.g., Card et al. 2016). This is particularly important in analyses of gender pay gaps, as the available evidence shows that these gaps can vary considerably across organizations (e.g., Heinze and Wolf 2010).

3 Institutional background

Finnish universities and the pay system for academic employees

The university system in Finland consists of ten multidisciplinary universities, two universities of technology, one university of the arts and one independent business school[1]. The universities are administered by the state, and the majority of their funding comes from the state budget and other public sources. In recent years, and following international patterns (Vincent-Lancrin 2009), the allocation of public funding has become more closely tied to university-specific output; the level of state funding is mainly based on universities' teaching loads (the number of graduates and course credits) and research achievements (the number and quality of publications) and, to a lesser extent, on university-specific and strategic factors. Furthermore, project-based research funding from external sources has become an increasingly important component of university budgets over the past decades.

In 2014, there were 17,653 researchers and university instructors in Finnish universities (AFIEE 2014). Compared to other EU-27 countries, women are well represented in Finnish academia, with the share of females exceeding those of other countries at every level of the academic hierarchy (Figure 1); in 2012, 52% of the faculty was female in academic ranks typically held by recent PhD graduates (grade C) and more senior researchers (grade B). However, female researchers also seem to be underrepresented in top academic positions in Finland: the share of female professors (grade A), although high compared to other nations, was only 24%.

The university analyzed in this paper, the University of Jyvaskyla, is the sixth largest university in Finland based on student enrollment. The university includes seven faculties, each with a number of schools and disciplines: 1) education, 2) humanities, 3) information technology, 4) mathematics and science, 5) social sciences, 6) sport and health sciences and 7) business and economics. As illustrated in Table 1, the student and personnel characteristics of this university are comparable to those of other Finnish multidisciplinary universities. A distinguishing feature of the University of Jyvaskyla is the high representation of women in top academic ranks: the share of female professors (37%) exceeds the average of other universities (31%), partly reflecting differences in the disciplinary composition of Finnish universities.

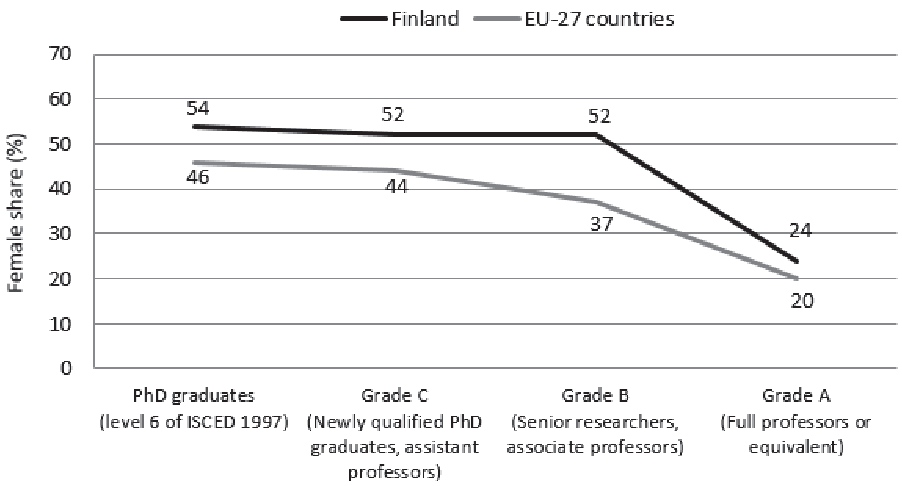
Academic earnings are set by a collective bargaining agreement, which applies to all university employees. The pay system is uniform across all universities and relates remuneration to the complexity of job tasks and personal performance by decomposing monthly earnings into two main components, namely, a task-specific component and a performance component (see Table A1

---

1 In addition to these PhD-granting research universities, the Finnish higher education system includes polytechnics (also referred to as universities of applied sciences) that specialize in tertiary level vocational education.

in the Appendix)[2]. The task-specific component is based on job complexity (measured on 11 levels) and determines the minimum earnings level. The performance component is proportional to task-specific component, varying from 0 to 46 percent depending on the employee's performance level (of 9 different levels). Additionally, employees can earn bonuses for supplementary assignments, such as administrative duties. In 2014, the average shares of the task-specific, performance and bonus components among full-time faculty members of Finnish universities were 79%, 19% and 2%, respectively, of total monthly earnings (AFIEE 2014).

FIGURE 1 Proportion of females at different academic career stages



Notes: Data source: "*She Figures 2012: Gender in Research and Innovation*", European Commission, Figure 2.1 and Table 3.1. Data for EU-27 countries were estimated.

2 For an extensive description of the pay system, see the "General collective agreement for universities" (downloadable at www.sivistystyonantajat.fi/tiedostopankki/158, viewed 24 July 2017).

TABLE 1 Comparison of universities in 2014

	Other multidisciplinary universities (average values)	University of Jyvaskyla
Students [a]		
Number of students (bachelor's and master's levels)	10 333	11 613
Number of graduates (master's degree)	1 103	1 452
Number of graduates (PhD)	129	158
All faculty members		
Employees [a]	1 254	1 409
Share of female employees (%)	47	49
Earnings (euros)	3 720 [b]	3703 [c]
Gender earnings gap (%)	−9.7 [b]	−9.6 [c]
Job complexity level (min = 1, max = 11)	4.7 [b]	4.9 [c]
Journal articles/faculty members	0.95	0.89
Full professors		
Share of faculty members (%)	17	16
Female professors (%)	31	37
Earnings (euros)	6837	6831
Gender earnings gap (%)	−3.2	−3.8
Job complexity level (min = 1, max = 11)	8.7	8.5

Notes: [a] Student and employee figures exclude the University of Helsinki (which had 28 185 students and 3 941 faculty members); [b] Average values of earnings and job complexity levels are calculated for faculty members of all Finnish universities (including the University of Jyvaskyla) in 2012; [c] Authors' own calculation from the personnel data (2012). Data sources: Vipunen-database (The Finnish National Board of Education), AFIEE (2014), the Finnish Union of University Professors.

Job complexity ladder, employee evaluations and promotions

When appointed to a university, a new employee typically starts a fixed term of employment lasting up to 5 years. After holding a temporary research or teaching position, the employee may be considered for a permanent appointment (an employment contract of indefinite duration), subject to satisfactory job performance. At the time of recruitment, the employee is assigned to one of 11 job complexity levels, with higher complexity levels being associated with a wider variety of academic duties, more complex job tasks and greater responsibility. There is a built-in relationship between the job complexity ladder and the hierarchy of occupations, as illustrated in Table 2: early career researchers, such as PhD students and teaching assistants, typically work at complexity levels 1–4, lecturers and researchers with more seniority at levels 5–7 and full professors at levels 8–11. Two details from Table 2 should be emphasized. First, each occupation has its own job complexity ladder. For example, within the rank of full professor, there exists a four-step ladder with job complexity levels ranging from 8 to 11. Second, job complexity levels overlap occupations; for example, senior researchers with the longest tenures may reach job complexity level 8, which is the typical starting level for newly hired full professors.

TABLE 2 Job complexity levels and occupations

	Job complexity level			
Occupation	Typical	Mode	Mean	Std. Dev.
Doctoral student	1–4	2	2.4	0.9
Researcher	2–5	3	3.5	1.3
Teaching assistant	2–5	3	3.7	1.0
University instructor	4–5	4	4.3	0.5
Postdoctoral researcher	5–6	5	5.2	0.4
Lecturer	5–7	6	5.8	0.9
Senior assistant	5–7	6	5.9	0.6
Senior researcher	5–8	6	6.1	0.9
Full professor	8–11	8	8.5	0.8

Notes: Authors' own calculations from the personnel data used in the following analysis.

Job complexity and performance levels are evaluated independently in an assessment meeting between a supervisor and an employee[3]. The assessment meeting is typically held once every two years, but the employee is entitled to request a re-assessment in the event of significant changes in his or her job duties. The job complexity level is assessed based on a job description, which includes all the essential duties and responsibilities of the employee. The assessment of personal performance is based on three different criteria: (1) teaching merit, (2) research achievements and (3) societal engagement and contributions to the university community. Each of these criteria is rated on a nine-point scale ranging from "very low" to "excellent" based on a performance evaluation of the assigned tasks and duties. The overall performance rate is obtained as a weighted sum of rates on different criteria, weighted by the share of working time devoted to each activity. After the job complexity and performance evaluations are agreed upon by the employee and the supervisor, the central university administration appraises the performance evaluations to ensure that performance is assessed consistently across employees in the same discipline, occupation and job complexity level.

[3] At the time of the recruitment, each employee is assigned a supervisor, typically the head or deputy head of a department.

A promotion on a job complexity ladder – i.e., an increase in an employee's job complexity level – is always associated with an increase in the variety in, complexity of and responsibility associated with the employee's job duties. Moreover, because the job complexity level essentially determines the minimum earnings level of an employee (see Table A1), a promotion on the complexity ladder is always accompanied by a pay increase. At the lowest job complexity levels, 1–4, promotion is typically the result of progress in PhD studies and increased teaching responsibilities. At higher complexity levels, 5–11, promotion involves a diversification of academic tasks (e.g., research, teaching, administrative duties, thesis supervision) and more responsibility and job complexity (e.g., heavier teaching loads, teaching more advanced courses, managing research projects, serving as the vice-head or head of a department).

An employee can be promoted on a job complexity ladder in three different ways. First, the job complexity level may be increased during an assessment meeting with a supervisor, which is organized biennially without the need for an employee request. Second, the employee can request a reassessment of job complexity if he or she is unsatisfied with the current assessment (e.g., due to the notable changes in his or her job duties and responsibilities after the previous assessment meeting). Third, the employee can apply and be appointed to a new occupation higher on the job complexity ladder.

4 Determinants of promotions and earnings

4.1 Data and empirical approach

The data employed are drawn from the personnel records of a Finnish university for the 2006–2012 period. This panel data include all full-time faculty members, with 8894 observations on 2583 individuals. The data set contains the following information for each individual[4]: personal id number, observation year, monthly earnings, gender, age, tenure, highest degree, department, occupation (academic rank), job complexity level, personal performance level and annual number of variously classified publications. Our data differ from those of earlier studies in two important ways. First, the data are well balanced by gender, with a proportion of women of approximately 48 percent. Second, the panel structure of the data allows us to track individuals over time; with few exceptions (Binder et al. 2010; Bratsberg et al. 2010; Haeck and Verboven 2012), the majority of the previous research on academic pay gaps has relied on cross-sectional data.

Table 3 summarizes the personnel data, showing that the average monthly earnings of female researchers were approximately 12% lower than those of their male colleagues. The mean values of background characteristics indicate that, on average, female faculty members were younger, had shorter tenures,

[4] These variables are described in more detail in the Appendix.

were less likely to hold a doctoral degree and worked at lower job complexity and performance levels than male faculty. The distributions of employees by occupation show that women were significantly more likely to work as university instructors and less likely to work as full professors than men. Furthermore, compared to men, women published a lower number of peer-reviewed international articles. A more detailed examination of the data implies that the lower publication activity of female researchers is partly explained by the concentration of women in disciplines (departments) with lower average output of international articles. In the following analysis, we will examine gender differences in research productivity more thoroughly by estimating a set of research output models.

The last panel of Table 3 reports the yearly promotion rates by gender. The reported promotion rates – defined as the fraction of employees whose job complexity level increased in consecutive years – reveal that a higher fraction of women were promoted than men and that a major portion of promotions occurred at lower rungs of the job complexity ladder. The promotion rate for all employees was 12.6%. Approximately one-fourth (24.5%) of all promotions were accompanied by a change of occupation. The majority of promotions (78%) consisted of shifts to the next level on the job complexity ladder, with only 22% increasing by two job complexity levels.

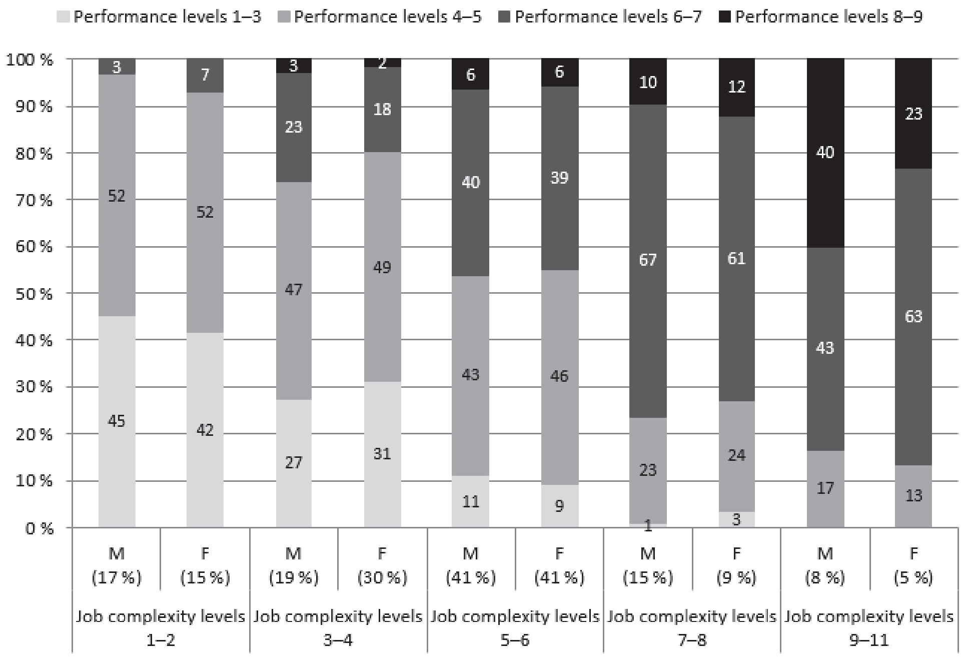
The joint distribution of employees' job complexity and performance levels in Figure 2 illustrates that a higher proportion of men than women were working at the highest job complexity levels; in 2012, 23% of men were working at complexity levels 7–11, compared to 14% of women. Furthermore, at the top of the job complexity ladder, men were likely to have higher performance levels than women: among those working at the highest job complexity levels (9–11), 40% of men and 23% of women attained the highest performance levels (8–9). The joint distribution also shows that job complexity and performance levels were positively related, indicating that the performance level was higher for those higher up the job complexity ladder (in 2012, the correlation coefficient between job complexity and performance level was 0.59). These observations raise two key questions that we address in this study: (1) Does the segregation of women at lower job complexity levels result from gender bias in promotion decisions and/or in entry-level job complexity levels? That is, do female employees encounter barriers to reaching higher levels on the job hierarchy? (2) Are the higher performance evaluations of men determined by actual gender differences in worker productivity or do they reflect undervaluation of female researchers' academic achievements?

TABLE 3 Descriptive statistics (2006–2012)

	Males	Females	Gender difference	t-test
	Mean (Std. Dev.)	Mean (Std. Dev.)		
Monthly earnings (euros)	3701 (1521)	3257 (1204)	+444	***
Job complexity level	5.1 (2.3)	4.6 (2.0)	+0.5	***
Performance level	5.5 (1.5)	5.1 (1.4)	+0.4	***
Age (years)	41.3 (11.5)	40.5 (10.8)	+0.8	***
Tenure (years)	8.5 (9.0)	7.0 (7.6)	+1.5	***
Education (%)				
Master's degree or lower	32.8	43.5	−10.7	***
Licentiate's degree	6.6	7.6	−1.0	*
Doctoral degree (PhD)	60.6	48.9	+11.7	***
Publication counts (per year)				
International refereed publications	1.8 (3.3)	0.8 (1.8)	+1.0	***
National refereed publications	0.2 (0.9)	0.3 (0.8)	−0.1	*
Other publications	1.0 (2.1)	0.7 (1.6)	+0.3	***
Occupation (%)				
Doctoral student	20.1	24.4	−4.3	***
Teaching assistant	4.7	7.4	−2.7	***
Researcher	13.4	12.1	+1.3	*
University instructor	3.3	10.4	−7.1	***
Postdoctoral researcher	9.5	7.9	+1.6	**
Senior assistant	6.2	4.5	+1.7	***
Senior researcher	6.9	5.0	+1.9	***
Lecturer	15.7	17.1	−1.4	*
Full Professor	18.7	9.2	+9.5	***
Other occupation	1.5	2.0	−0.5	
Promotion rates (%)[a]				
Promotion rate	11.9	13.4	−1.5	*
Promotion rate, job complexity level ≤ 4	25.1	23.6	+1.5	
Promotion rate, job complexity level > 4	5.0	5.2	−0.2	
Observations	4674 (52.6 %)	4220 (47.4 %)		

Notes: [a] Promotion = increase in job complexity level in consecutive years. Reported promotion rates are averages of yearly promotion rates. Rates are based on employees who worked for (at least) two consecutive years. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

FIGURE 2 Job complexity and performance levels in 2012



Using our longitudinal personnel data, we first evaluate the role of gender in promotion decisions by running a linear probability model on whether the employee's job complexity level increased between two periods. Second, we estimate ordered probit models of job complexity and performance levels to analyze gender bias in assignment to different job ladders and in employee performance evaluations. Third, we assess the robustness of gender gaps in total earnings by estimating a set of earnings equations using standard OLS regressions. Finally, we conduct additional analyses to determine (1) whether gender differences in the production of peer-reviewed articles exist, (2) whether a gender difference in the probability of working as a full professor exists and (3) whether a gender pay gap exists within the full professor rank. In all estimated models, we control for an appropriate set of individual qualifications, job characteristics and research productivity variables.

We include age and job tenure to control for the employee's previous work experience: age acts as a proxy for potential total work experience, and job tenure measures the time that has passed since an employee became employed at the university. To account for the effects of education level, we include dummy variables for the highest degree completed. To allow for possible career outcome differences between academic disciplines, we employ dummy variables for departments as proxy variables. The discipline controls are particularly important for the earnings equations because disciplines may differ significantly with respect to outside wage offers, and academic earnings may be inversely related to the proportion of women in the discipline (Bellas 1997; Umbach 2006).

In the earnings equations, we include a dummy variable for administrative duties for two reasons: to account for the additional compensation received for performing these duties and the time spent on these duties. Remuneration for administrative tasks yields an additional source of gender bias in earnings if men are more likely to be assigned to administrative positions. To assess potential gender bias in the assignment of administrative tasks, we regressed a dummy variable for these duties on a gender dummy and set of individual background variables (see Table A2 in the Appendix); the gender coefficients were consistently close to zero and statistically insignificant, implying that gender did not play a role in the assignment of administrative duties.

To evaluate the role of worker-specific productivity on earnings and promotions, we include three research productivity variables in our models: the number of peer-reviewed international and national articles and other publications (e.g., books, book chapters, working papers). The distinction among publication types is essential, as some academic disciplines primarily focus on international publications, whereas others also emphasize the importance of national publications; moreover, academic disciplines differ with respect to how they weight journal articles relative to other publications (Räty and Bondas 2008).

As our data only include the annual publication records of employees working at this university (i.e., we have no information on past research achievements or merit beyond this university), we use the contemporary publication count as a proxy variable for employees' past research productivity. Hence, we implicitly assume that individual research productivity is relatively stable over time. Because this assumption may fail to hold in practice, we also use the cumulative publication count in the previous periods to assess the effects of past productivity on earnings and promotions. Furthermore, because earnings and promotion decisions may depend not only on absolute worker output but also on relative output, we also employ relative publications – calculated by dividing the publication count for worker i during a given period by the average number of publications in worker i's discipline (department) during that period – in our analysis.

In the absence of information on the quality of individual research output, we cannot directly analyze whether higher quality research was rewarded with higher earnings and/or promotions. However, the distinction among publication types provides an indirect way to assess the role of research quality in earnings and promotions: international peer-reviewed articles are likely to carry more weight in performance evaluations than other publications. Hence, we expect to observe larger positive coefficients on these articles than on other publication variables in the estimated earnings and promotion equations.

4.2 Determinants of promotions

To analyze the determinants of promotions, we estimate the following linear probability model:

$$Promoted_{ijdt} = a + \beta Female_i + \boldsymbol{\gamma Publications_{it}} + \boldsymbol{\delta X_{i,\,t-2}} + \pi_{j,\,t-2} + \theta_d + \varepsilon_{ijdt} \qquad (1)$$

where the dependent variable $Promoted_{ijdt}$ is a dummy variable that equals one if employee i's job complexity level increased from the previous period. This is an appropriate definition of a promotion because, as described above, higher complexity levels entail more demanding job tasks and greater responsibilities. Because employees typically have a promotion opportunity once every two years, the dependent variable of the base specifications is equal to one if an employee was promoted between year $t-2$ and year t and zero otherwise. The dummy variable $Female_i$ equals one if employee i is a female. $\boldsymbol{Publications_{it}}$ is a vector of three publication variables (peer-reviewed international articles, national articles and other publications) that indicate the sum of employee i's publications over the two previous years. $\boldsymbol{X_{i,\,t-2}}$ is a vector of two-year-lagged control variables, including age, tenure, dummy variables for education levels and a dummy variable indicating whether an employee's education level changed from the previous period (i.e., between $t-2$ and t) and θ_d are department dummies. Furthermore, control variables include the job complexity level in the previous period, $\pi_{j,\,t-2}$, to account for the fact that there are a limited number of job complexity levels and that promotion probabilities may differ across job complexity levels (see Table 3)[5].

Table 4 presents the main results of the linear probability models of promotions within job complexity levels. The estimated gender coefficients in columns 1 and 2 indicate that the probability of promotion was lower for women only before controlling for research productivity differences. The estimates in the next columns show that this conclusion holds after controlling for the background characteristics of a worker (column 3), after using relative publications instead of publication counts (column 4) and after redefining the dependent variable to account for year-to-year promotions (column 5). Hence, the results suggest that female researchers were as likely to be promoted as their similarly productive and qualified male colleagues.

Furthermore, consistent with previous studies (Ward 2001; Ginther and Hayes 2003), the results in Table 4 suggest that higher research productivity was associated with a higher probability of being promoted, with the coefficient estimates in columns 2, 3 and 5 indicating that the promotion probability increased with the publication count. These estimates imply that national articles carried more weight in promotion decisions than did international articles. This finding is explained by the sensitivity of the results to the extensive publication records of a few researchers; for example, excluding the top 5% of observations

---

5 The job complexity level in the previous period is entered as a linear term; using dummy variables instead did not change the main results.

for all peer-reviewed articles more than doubled the coefficient estimate for international articles (to 0.033) in the model in column 3, while the other publication coefficients remained nearly unchanged. Hence, the number of internationally published, peer-reviewed articles seems to have been the primary research output measure in promotion decisions. This conclusion is confirmed by the research productivity estimates presented in column 4, which suggest that both absolute and relative research output were factors in promotion decisions: the relative number of peer-reviewed international articles was positively and significantly related to promotion prospects, whereas the relative output of peer-reviewed national articles was not a significant factor in promotion decisions.

The results in Table 4 strongly indicate that, conditional on individual research output, gender was not a determinant of promotions. In addition to differences in promotions, a gender-biased job hierarchy may result not only from gender differences in promotion probabilities but also from gender differences in the assignment of employees to job levels upon hiring. To examine the role of gender in the determination of job levels more carefully, we estimate an ordered probit model of the job complexity level. In other words, we estimate a latent variable model of the following form[6]:

$$\textit{Job complexity level}_{idt}^{*} = \beta \textit{Female}_i + \gamma \boldsymbol{\mathit{Publications}_{it}} + \delta X_{it} + \theta_d + \varepsilon_{idt} \qquad (2)$$

where the latent unobserved variable *Job complexity level*$_{idt}$* takes values in {1, 2, 3, … , 11}. The dummy variable $Female_i$ equals one if employee i is a female; $\boldsymbol{\mathit{Publications}_{it}}$ is a vector of contemporaneous publication counts of peer-reviewed international and national articles and other publications; $\boldsymbol{X_{it}}$ is a vector of control variables, including age and tenure (as well as their squared terms) and dummy variables for the education level; and $\boldsymbol{\theta_d}$ are department dummies. The parameters of the model (β, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$, θ) are estimated by maximum likelihood estimation.

Table 5 presents the coefficients of the ordered probit models of the job complexity level[7]. According to the results, there is weak or no evidence that gender plays a role in the assignment of employees to a level on the job hierarchy: the estimated gender coefficient is statistically significant when contemporaneous publication counts are used to account for research productivity differences (column 1) but statistically insignificant when past publications are used to measure worker output (columns 2–3). The coefficient estimates on the publication variables indicate that more productive (both in absolute and relative terms) faculty members were more likely to work at higher levels of the job complexity ladder.

---

6 See Wooldridge (2010, pp. 655–657) for a derivation of the ordered probit model from a latent variable model.

7 The marginal effects of the gender variable are reported in the upper panel of Table A1 in the Appendix.

TABLE 4 Determinants of promotion (linear probability models of job complexity level)

Dependent dummy variable = 1 if job complexity level increased from previous period					
	Promoted between				
	$t-2$ and t				$t-1$ and t
Female	−0.031 (0.016)*	−0.001 (0.015)	0.009 (0.016)	0.003 (0.016)	0.008 (0.009)
Publications		*Count*	*Count*	*Relative*	*Count*
International refereed		0.015 (0.002)***	0.013 (0.002)***	0.029 (0.005)***	0.012 (0.002)***
National refereed		0.024 (0.007)***	0.022 (0.006)***	−0.001 (0.001)	0.015 (0.005)***
Other publications		0.009 (0.002)***	0.009 (0.002)***	0.012 (0.003)***	0.005 (0.002)***
Controls (period)	($t-2$)	($t-2$)	($t-2$)	($t-2$)	($t-1$)
Worker characteristics	No	No	Yes	Yes	Yes
Department dummies	No	No	Yes	Yes	Yes
Job complexity level	Yes	Yes	Yes	Yes	Yes
Sample: Those who had worked for at least	*Three consecutive years*				*Two consecutive years*
Publication variables: Publications for the	*Two previous years*				*Previous year*
Period	2008–2012				2007–2012
$R^2_{adjusted}$	0.18	0.22	0.27	0.26	0.13
Observations	3782	3782	3695	3634	5788

Notes: Cluster-robust standard errors in parentheses (clustered at the worker level). All models include a constant term. Worker characteristics include age, tenure, dummy variables for education levels and a dummy variable indicating whether an employee's education level changed from the previous period. When the models in the last three columns were estimated with department dummy variables as explanatory variables, the coefficient estimates were virtually unchanged. Using probit or logit models instead of linear probability model produced qualitatively similar results. Full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

TABLE 5 Ordered probit model of job complexity level

	(1)	(2)	(3)
Female	−0.114	−0.069	−0.089
	(0.044)***	(0.063)	(0.063)
Publications	*Count*	*Count*	*Relative*
International refereed	0.113	0.084	0.159
	(0.009)***	(0.008)***	(0.026)***
National refereed	0.062	0.072	0.005
	(0.020)***	(0.023)***	(0.007)
Other publications	0.070	0.054	0.077
	(0.009)***	(0.007)***	(0.012)***
Controls			
Worker characteristics	Yes	Yes	Yes
Department dummies	Yes	Yes	Yes
Sample	*Full sample*	*Those who had worked for at least three consecutive years*	
Publication variables: Publications for the	*Contemporaneous year*	*Two previous years*	
Period	2006–2012	2008–2012	
Observations	8894	3782	3744
Pseudo R^2	0.31	0.33	0.31
Log pseudolikelihood	−12907	−5225	−5289

Notes: Table reports the ordered probit coefficients. Cluster-robust standard errors in parentheses (clustered at the worker level). Worker characteristics include age, age^2, tenure, tenure^2 and dummy variables for education levels. The full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

4.3 Determinants of performance evaluations

To examine the determinants of performance evaluations, we employ an ordered probit model similar to that in equation (2), with the individual performance level now used as the dependent variable. Columns 1–3 in Table 6 present the main results of the ordered probit analysis[8]. All reported models condition on a set of worker background characteristics, research productivity and job complexity. Controlling for the job complexity level is important because

[8] The table reports the ordered probit coefficients. The marginal effects of the gender variable are reported in the lower panel of Table A1 in the Appendix.

performance levels are positively related to job complexity levels (see Figure 2); however, excluding job complexity from the models does not alter the conclusions of the analysis presented here. The estimates in column 1 are conditioned on contemporaneous publication counts to control for individual research output. The estimated gender coefficient is negative and statistically significant at the 10% level. The significant gender difference disappears when past publication output is used to control for research productivity in columns 2 and 3, suggesting that male and female employees with similar qualifications and research outputs received similar performance evaluations. Hence, in contrast to some previous studies (Bartol 1999; Castilla 2012), our results suggest that gender plays a negligible or no role in employee performance evaluations. The estimated coefficients on the publication variables indicate a positive relationship between actual output and assessed performance, implying that better-performing employees – whether measured in absolute or in relative terms – are likely to receive better performance evaluations.

Furthermore, to examine whether gender differences in changes in performance evaluations exist, we estimate a linear probability model that parallels that in equation (1), with the dependent variable now being a dummy variable that equals one if the employee's performance level increased between year $t-2$ and year t and zero otherwise[9]. To account for the fact that performance level changes among employees who moved up or down the job complexity ladder were likely less related to job performance than to other considerations (e.g., promotions were often associated with a decrease in individual performance), the estimated model is based on a sample of employees whose job complexity remained unchanged from the previous period. The results of the linear probability models of performance level increments are presented in columns 4 and 5 of Table 6. The nonsignificant coefficient estimates on the female variable indicate that the probability of being upgraded to a higher performance level did not depend on gender. The coefficients of the publication variables show that employees who produced more peer-reviewed articles were more likely to be upgraded to a higher performance level, while output of other publications had no effect on the probability of being upgraded.

[9] Linear probability models for year-to-year increments of performance level provided qualitatively similar results.

TABLE 6 Determinants of individual performance level

	Ordered probit models of performance level			Linear probability models: level increased between $t-2$ and t	
Female	−0.084 (0.044)*	−0.028 (0.064)	−0.035 (0.064)	−0.026 (0.021)	−0.029 (0.021)
Job complexity level	0.077 (0.017)***	0.055 (0.026)**	0.076 (0.025)***		
<u>Publications</u>	*Count*	*Count*	*Relative*	*Count*	*Relative*
International refereed	0.040 (0.008)***	0.026 (0.007)***	0.062 (0.015)***	0.011 (0.002)***	0.027 (0.007)***
National refereed	0.044 (0.018)**	0.045 (0.020)**	0.005 (0.002)**	0.016 (0.007)**	0.003 (0.001)***
Other publications	0.050 (0.008)***	0.043 (0.007)***	0.055 (0.011)***	0.002 (0.003)	0.002 (0.004)
<u>Controls (period)</u>	(t)	(t)	(t)	($t-2$)	($t-2$)
Worker characteristics	Yes	Yes	Yes	Yes	Yes
Department dummies	Yes	Yes	Yes	Yes	Yes
Sample	*Full sample*	*Those who had worked for at least three consecutive years*		*Those who had worked for at least three consecutive years and whose job complexity levels remained unchanged between t−2 and t*	
Publication variables: Publications for the	*Contemp. year*	*Two previous years*		*Two previous years*	
Period	2006–2012	2008–2012		2008–2012	
Observations	8891	3780	3742	2767	2732
Pseudo R^2	0.13	0.13	0.13		
$R^2_{adjusted}$				0.10	0.09
Log pseudolikelihood	−13904	−5750	−5707		

Notes: The first three columns report the coefficients of the ordered probit models. Cluster-robust standard errors in parentheses (clustered at the worker level). The dependent dummy variable of the linear probability models equal one if employee's performance level increased between $t-2$ and t and zero otherwise. Worker characteristics include age, age^2, tenure, tenure^2 and dummy variables for education levels. Both linear probability models include a constant term. The full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

4.4 Earnings differentials

In order to assess the magnitude of the gender pay gap, we estimate the following earnings equation using standard OLS regression:

$$log(Earnings_{ijdt}) = a + \beta Female_i + \boldsymbol{\gamma Publications_{it}} + \boldsymbol{\delta X_{it}} + \theta_d + \lambda_t + \varepsilon_{ijdt} \quad (3)$$

where the dependent variable is a logarithm of monthly earnings (in euros) for employee *i* in department *d* at job complexity level *j* in year *t*; $Female_i$ equals one if an employee is a female; $\boldsymbol{Publications_{it}}$ is a vector of contemporaneous publication counts of peer-reviewed international and national articles and other publications; $\boldsymbol{X_{it}}$ is a vector of control variables, including age and tenure (as well as their squared terms), a dummy variable for administrative duties and dummy variables for the education level; and θ_d are department dummies and λ_t are year dummies. If there exists a gender gap in earnings, we would expect to observe a statistically significant nonzero value for the coefficient of the *Female* variable, β.

Table 7 summarizes the main results of the earnings equations. According to the gender coefficient in column 1, female researchers earned approximately 11% less than their male co-workers. The estimates in columns 2 and 3 illustrate that the male premium in earnings is mainly explained by differences in research productivity and individual background characteristics: after adjusting for these differences, a gender gap of approximately 2% remains. Adding controls for occupations in column 4 significantly improves the fit of the model and yields a statistically insignificant gender gap of approximately 1%, suggesting no gender pay gaps among occupations. If the assignment of employees to different occupational levels depended on gender, then the inclusion of occupation dummies would bias the estimated gender earnings gap downward; however, as the analysis above suggests, gender was not a significant determinant of position on the job hierarchy in this particular organization. Using past research output instead of contemporaneous output to control for research productivity in column 5 produces a statistically insignificant gender earnings gap of approximately 1%. The earnings equations reported in the table only control for publication counts and, hence, do not account for differences in relative research productivity. However, re-estimating the models in columns 3–5 using relative publications variables instead of publication counts lead to similar results regarding the gender pay gap: the gender coefficient is close to zero and typically statistically insignificant.

The previous results provide strong evidence that the observed gender gap in average earnings is mainly attributable to worker differences in background characteristics and research productivity. To further examine whether there were gender differences in *pay changes*, we replaced the dependent variable of the earnings equation (3), the logarithm of earnings, with the difference in logarithmic earnings between two consecutive years and used one-year lagged publication and background variables instead of contemporaneous variables as regressors. The results of the estimated pay-change equation are reported in column 6. The results indicate that gender was not a determinant of year-to-year earnings changes. The coefficient estimates on the publication variables show that earnings changes were positively related to research output, implying that increments in earnings were higher for more productive workers.

Data limitations prevent us from directly assessing the robustness of our results to the inclusion of productivity measures other than research output. One important measure might be the amount of time devoted to teaching. In the absence of teaching data, we evaluated the sensitivity of the gender and publication coefficients to the omission of teaching load variables by re-estimating the model in column 3 after excluding the most teaching-intensive occupations, namely, university instructors and lecturers. The resulting coefficient estimates on the gender and publication variables were virtually unaffected, implying that the main results are not altered by the omission of teaching load variables[10].

[10] The results of these estimations are available upon request.

TABLE 7 Earnings equations

	Dependent variable					
	log(earnings)					Difference in log(earnings) between *t* and *t−1*
Female	−0.110 (0.016)***	−0.063 (0.014)***	−0.024 (0.007)***	−0.008 (0.005)	−0.011 (0.010)	0.0006 (0.0017)
<u>Publication counts</u>						
International refereed		0.037 (0.003)***	0.020 (0.002)***	0.010 (0.001)***	0.013 (0.001)***	0.0010 (0.0003)***
National refereed		0.055 (0.009)***	0.014 (0.003)***	0.007 (0.002)***	0.015 (0.004)***	0.0021 (0.0010)**
Other publications		0.043 (0.004)***	0.012 (0.002)***	0.006 (0.001)***	0.009 (0.001)***	0.0011 (0.0004)**
<u>Controls (period)</u>	(*t*)	(*t*)	(*t*)	(*t*)	(*t*)	(*t−1*)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Worker characteristics	No	No	Yes	Yes	Yes	Yes
Department dummies	No	No	Yes	Yes	Yes	Yes
Occupation dummies	No	No	No	Yes	No	No
Sample	*Full sample*				*Those who had worked for at least three consecutive years*	*Those who had worked for at least two consecutive years*
Publication variables: Publications for the	*Contemporaneous year*				*Two previous years*	*Previous year*
Period	2006–2012				2008–2012	2007–2012
$R^2_{adjusted}$	0.03	0.20	0.79	0.90	0.80	0.05
Observations	8894	8894	8894	8894	3782	5788

Notes: Earnings = monthly earnings in euros. Cluster-robust standard errors in parentheses (clustered at the worker level). All models include a constant term. Worker characteristics include age, age^2, tenure, tenure^2, a dummy variable for administrative duties and dummy variables for education levels. (The linear probability model in the last column also includes a dummy variable indicating whether an employee's education level changed from the previous period as a control variable.) The full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

4.5 Additional findings

Gender differences in research productivity

The above results suggest that gender gaps in promotion rates and earnings partly reflect gender differences in research productivity. The lower research productivity of female academics is widely acknowledged in the literature (Schneider 1998; Xie and Shauman 1998). Other empirical studies suggest that gender gap in research output cannot be fully explained by differences in researcher characteristics (e.g., experience and academic rank) or by other factors, such as the concentration of female researchers in academic disciplines with less publishing (e.g., Toutkoushian and Bellas 1999; Hesli and Lee 2011).

To analyze gender differences in research productivity in more detail, we estimate the following equation using OLS, Poisson and negative binomial regression models:

$$Articles_{ijdt} = a + \beta Female_i + \gamma Other\ publications_{it} + \boldsymbol{\delta X_{it}} + \pi_{jt} + \theta_d + \varepsilon_{ijdt} \qquad (4)$$

where the dependent variable is the annual number of peer-reviewed articles (both international and national) of employee i at job complexity level j in department d in year t; $Female_i$ equals one if an employee is a female; $Other\ publications_{it}$ is the annual number of other (non-refereed) publications; $\boldsymbol{X_{it}}$ is a vector of control variables, including age and tenure (as well as their squared terms), a dummy variable for administrative duties and a dummy variable for a doctoral degree; and π_{jt} are job complexity level dummies and θ_d are department dummies.

The regression results of equation (4) are presented in Table 8. The gender coefficient from the OLS regression in column 1 indicates that female researchers produced, on average, approximately one fewer article than their male colleagues. The results in column 2 suggest that this gender gap in research productivity is mainly attributable to differences in worker characteristics, and the gender coefficient is no longer statistically significant when these differences are accounted for. However, because the OLS regression assumes a continuous dependent variable and is therefore not appropriate for the analysis of count dependent variable, we also estimated research output using methods designed for count data, namely, Poisson regression (column 3) and negative binomial regression (column 4). The gender coefficients from these preferred regressions indicate that female researchers produced statistically significantly fewer peer-reviewed articles than male researchers with similar background characteristics.

The existing literature proposes several potential explanations for the gender gap in research output. First, female researchers' research output might be adversely affected by childbearing and heavier engagement in childcare and other household responsibilities (e.g., Stack 2004). Second, female faculty members may use more of their working time to activities other than research (Toutkoushian and Bellas 1999; Link et al. 2008), possibly due to their stronger preferences for or motivation to engage in non-research activities (e.g., Bentley and Kyvik 2013). Third, insufficient resources and weaker research networks may diminish the publication output of female researchers, especially in male-dominated disciplines if researchers tend to co-author with colleagues of the same sex (McDowell and Smith 1992) and if research productivity increases with co-authorship (Hollis 2001). Fourth, journal editors and reviewers may discriminate against female authors, leading to higher rejection rates for female-authored manuscripts (Ferber and Teiman 1980). However, the empirical support for these explanations is ambiguous, as studies have shown that (1) female researchers with dependent children have similar research productivity as male researchers (e.g., Sax et al. 2002), that (2) additional time spent on other activities – most notably, teaching – does not have a negative effect on research output (Shin and Cummings 2010) and gender is a weak predictor of research time amongst university faculty (Bentley and Kyvik 2013) and that (3) gender does not play a role in the article review process (Abrevaya and Hamermesh 2012).

The estimated coefficients of other covariates also reveal some interesting relationships. First, research output was lower for older workers and increased with tenure but at a diminishing rate. Second, the coefficients imply that time spent on administrative duties had a negative effect on research productivity. Third, those who had higher outputs of other publications (e.g., non-refereed book chapters, discussion papers) produced more peer-reviewed articles.

Female professors and the gender pay equity of professors

The results above show little or no evidence of gender bias in the assignment of faculty members to different levels on the job complexity ladder. However, a potential obstacle to career advancement was identified for female researchers: women may be less likely to achieve the full professor rank than men. As reported in Table 3, approximately 19% of men were working as full professors compared with 9% of women. To assess whether this difference was attributable to differences in worker characteristics, the first column of Table 9 reports the gender difference in the likelihood of holding a full professor position among employees working at job complexity levels 6–11. The female coefficient suggests that, conditional on worker and job characteristics, female employees were 6% less likely to work as professors than equally qualified males. Although women were underrepresented in professor positions, women who had achieved professorships earned equal pay for equal work: the results of the earnings equation in the second column of the table imply no gender pay gap within the professor rank.

TABLE 8 Determinants of research productivity

Dependent variable: Number of peer-reviewed articles in year *t*				
	OLS	OLS	Poisson regression	Negative binomial regression
Female	−1.032	−0.071	−0.108	−0.128
	(0.119)***	(0.080)	(0.057)*	(0.051)**
Age		0.103	−0.025	−0.028
		(0.043)**	(0.005)***	(0.004)***
Age^2		−0.0015		
		(0.0005)***		
Tenure		0.045	0.037	0.046
		(0.019)**	(0.011)***	(0.010)***
$Tenure^2$		−0.0016	−0.0013	−0.0015
		(0.0006)**	(0.0003)***	(0.0003)***
Doctoral degree		0.373	0.458	0.502
		(0.088)***	(0.080)***	(0.069)***
Administrative duties		−0.322	−0.155	−0.139
		(0.277)	(0.090)*	(0.080)*
Other publications		0.118	0.051	0.082
		(0.027)***	(0.009)***	(0.011)***
<u>Controls</u>				
Job complexity level	No	Yes	Yes	Yes
Department	No	Yes	Yes	Yes
$R^2_{adjusted}$	0.03	0.34		
Pseudo R^2			0.31	0.13
Observations	8894	8894	8894	8894
Log pseudolikelihood			−14 878	−12 695
α				0.903***

Notes: Cluster-robust standard errors in parentheses (clustered at the worker level). All models include a constant term. In the Poisson and negative binomial regressions, only statistically significant worker and job characteristics were included in the models (i.e., age2 was excluded). * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

TABLE 9 Likelihood of working as a professor and professor earnings

	Probability of working as a professor	Earnings of professors
	Dependent variable:	
	= 1 if professor, = 0 otherwise	log(Monthly earnings)
Female	−0.060	0.003
	(0.036)*	(0.011)
Controls		
Year	Yes	Yes
Worker characteristics	Yes	Yes
Research productivity	Yes	Yes
Department	Yes	Yes
Sample	Job complexity level > 5	All professors
$R^2_{adjusted}$	0.24	0.56
Observations	3271	1262

Notes: Cluster-robust standard errors in parentheses (clustered at the worker level). Worker characteristics include age, age², tenure, tenure² and dummy variables for education levels (earnings equation also includes a dummy variable for administrative duties). Research productivity variables include publication counts (peer-reviewed international articles, peer-reviewed national articles and other publications) in period *t*. Both models include a constant term. Full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

5 Conclusions

This study employs personnel data to evaluate the role of gender in internal promotion, employee performance evaluation and earnings determination. Using detailed information on the complexity rating of job tasks to identify promotions along the job hierarchy, we show that male and female researchers were equally likely to be promoted, conditional on individual research productivity. The findings demonstrate that worker-specific productivity differences may be a primary reason for gendered promotion rates. An analysis of the determinants of employee performance evaluations reveals that gender played a negligible or no role in evaluation decisions. The observed male premium in earnings was mainly attributable to individual differences in research productivity and background characteristics: adjusting for these differences reduced the gender earnings gap from approximately 11% to approximately 1–2%.

Moreover, once the full set of controls was included, the gender coefficient was no longer statistically significant. Additionally, the results demonstrate that female researchers had lower research output than their male colleagues, even after conditioning on a set of worker characteristics, including age, tenure and academic discipline. Finally, the results suggest that female and male professors were paid equally, although female employees were less likely to work as full professors than equally qualified men.

The results indicate that higher research productivity was related to higher probabilities of being promoted to or working on the highest job ladders (academic ranks). The findings also confirm that more productive researchers received more favorable performance evaluations than others with similar background characteristics, implying that the available worker output information was effectively employed in the assessment of employee performance and was therefore likely to reduce the subjectivity of the evaluation process and result in more objective performance evaluations.

Our analysis employed publication counts to measure individual productivity. Other productivity measures, such as the quality of research and teaching, can also contribute to pay and promotion decisions. Earlier studies illustrate that publication quality, as measured by the number of citations (Moore et al. 1998; Bratsberg et al. 2010) or by the number of articles in top-tier journals (e.g., Hilmer and Hilmer 2005), is positively related to academic salaries. Our findings also provide some evidence that the quality of research matters for career advancement decisions: peer-reviewed international articles carried more weight in pay and promotion decisions than peer-reviewed national articles or other publications.

Employees may also receive rewards for their teaching load and skill, in terms of higher earnings and promotion probabilities. However, given (1) the theoretical arguments for why incentives for research productivity may have increased in universities (Remler and Pema 2009) and (2) the empirical findings suggesting that universities have become more inclined to make hiring, promotion and remuneration decisions largely based on research without regard to other achievements (Laband and Tollison 2003; Remler and Pema 2009), teaching may play a constantly diminishing role in various career decisions. In fact, the findings of several recent studies suggest that heavier teaching loads are penalized with lower earnings (e.g., Graves et al. 2002; Umbach 2006; Binder et al. 2012). Teaching might be less relevant to pay and promotion decisions at the university analyzed in this paper for several reasons. First, teaching loads are typically uniform within occupations, and the results are robust to the exclusion of occupations with more variable teaching loads. Second, the assessment of teaching skill is difficult, especially because student evaluations of instructors are not collected. Finally, the university's funding is closely tied to the number of research publications, giving supervisors strong incentives to emphasize publications in pay and promotion decisions.

Two conclusions can be drawn from our findings about the role of worker productivity in career outcomes. First, both absolute and relative individual output may be important factors in determining promotions, earnings and performance evaluations. Second, using contemporaneous and past productivity measures yielded qualitatively very similar results for the effects of worker output on earnings and performance evaluations, suggesting that information on employees' concurrent productivity provides a valid proxy for their past productivity.

References

Abrevaya, J. and Hamermesh, D.S. 2012. Charity and favoritism in the field: Are female economists nicer (to each other)?. *The Review of Economics and Statistics* 94(1), 202–207.

AFIEE. 2014. *Tilastojulkaisu 2014: Yliopistot* (in Finnish). The Association of Finnish Independent Education Employers: Helsinki, Finland.

Arulampalam, W., Booth, A.L. and Bryan, M.L. 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review* 60(2), 163–186.

Baker, G., Gibbs, M. and Holmstrom, B. 1994. The wage policy of a firm. *The Quarterly Journal of Economics* 109(4), 921–955.

Barbezat, D.A. 1991. Updating estimates of male-female salary differentials in the academic labor market. *Economics Letters* 36(2), 191–195.

Bartel, A.P. 1995. Training, wage growth, and job performance: Evidence from a company database. *Journal of Labor Economics* 13(3), 401–425.

Bartol, K.M. 1999. Gender influences on performance evaluations. In: G.N. Powell (Ed.) *Handbook of Gender and Work*, pp. 165–178. Thousand Oaks, CA, US: SAGE Publications.

Bellas, M.L. 1997. Disciplinary differences in faculty salaries: Does gender bias play a role? *The Journal of Higher Education* 68(3), 299–321.

Bentley, P.J. and Kyvik, S. 2013. Individual differences in faculty research time allocations across 13 countries. *Research in Higher Education* 54(3), 329–348.

Binder, M., Chermak, J., Krause, K. and Thacher, J. 2012. The teaching penalty in higher education: evidence from a public research university. *Economics Letters* 117(1), 39–41.

Binder, M., Krause, K., Chermak, J., Thacher, J. and Gilroy, J. 2010. Same work, different pay? Evidence from a US public university. *Feminist Economics* 16(4), 105–135.

Blackaby, D., Booth, A.L. and Frank, J. 2005. Outside offers and the gender pay gap: empirical evidence from the UK academic labour market. *The Economic Journal* 115(501), F81–F107.

Blau, F.D. and Kahn, L.M. 2000. Gender differences in pay. *Journal of Economic Perspectives* 14(4), 75–99.

Booth, A.L. 2009. Gender and competition. *Labour Economics* 16(6), 599–606.

Booth, A.L., Francesconi, M. and Frank, J. 2003. A sticky floors model of promotion, pay, and gender. *European Economic Review* 47(2), 295–322.

Bratsberg, B., Ragan, J.F. and Warren, J.T. 2010. Does raiding explain the negative returns to faculty seniority? *Economic Inquiry* 48(3), 704–721.

Card, D., Cardoso, A.R. and Kline, P. 2016. Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics* 131(2), 633–686.

Castilla, E.J. 2012. Gender, race, and the new (merit-based) employment relationship. *Industrial Relations* 51(S1), 528–562.

De la Rica, S., Dolado, J.J. and Vegas, R. 2010. Performance pay and the gender wage gap: evidence from Spain. *IZA Discussion Paper Series* No. 5032.

Dohmen, T.J., Kriechel, B. and Pfann, G.A. 2004. Monkey bars and ladders: the importance of lateral and vertical job mobility in internal labor market careers. *Journal of Population Economics* 17(2), 193–228.

Dohmen, T., Lehmann, H.F. and Schaffer, M.E. 2014. Wage policies of a Russian firm and the financial crisis of 1998: evidence from personnel data, 1997 to 2002. *ILR Review* 67(2), 504–531.

Elvira, M.M. and Graham, M.E. 2002. Not just a formality: Pay system formalization and sex-related earnings effects. *Organization Science* 13(6), 601–617.

Ferber, M.A. and Teiman, M. 1980. Are women economists at a disadvantage in publishing journal articles?. *Eastern Economic Journal* 6(3–4), 189–193.

Flabbi, L. and Ichino, A. 2001. Productivity, seniority and wages: new evidence from personnel data. *Labour Economics* 8(3), 359–387.

Francesconi, M. 2001. Determinants and consequences of promotions in Britain. *Oxford Bulletin of Economics and Statistics* 63(3), 279–310.

Ginther, D.K. and Hayes, K.J. 2003. Gender differences in salary and promotion for faculty in the humanities 1977–95. *The Journal of Human Resources* 38(1), 34–73.

Graves, P.E., Marchand, J.R. and Sexton, R.L. 2002. Hedonic wage equations for higher education faculty. *Economics of Education Review* 21(5), 491–496.

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

Heinze, A. and Wolf, E. 2010. The intra-firm gender wage gap: a new view on wage differentials based on linked employer-employee data. *Journal of Population Economics* 23(3), 851–879.

Hesli, V.L. and Lee, J.M. 2011. Faculty research productivity: Why do some of our colleagues publish more than others?. *PS: Political Science and Politics* 44(2), 393–408.

Hilmer, C.E. and Hilmer, M.J. 2005. How do journal quality, co-authorship, and author order affect agricultural economists' salaries? *American Journal of Agricultural Economics* 87(2), 509–523.

Hollis, A. 2001. Co-authorship and the output of academic economists. *Labour Economics* 8(4), 503–530.

Jawahar, J. and Williams, C.R. 1997. Where all the children are above average: a meta analysis of the performance appraisal purpose affect. *Personnel Psychology* 50(4), 905–925.

Jirjahn, U. and Stephan, G. 2004. Gender, piece rates and wages: evidence from matched employer-employee data. *Cambridge Journal of Economics* 28(5), 683–704.

Kangasniemi, M. and Kauhanen, A. 2013. Performance-related pay and gender wage differences. *Applied Economics* 45(36), 5131–5143.

Kelchtermans, S. and Veugelers, R. 2013. Top research productivity and its persistence: gender as a double-edged sword. *The Review of Economics and Statistics* 95(1), 273–285.

Kunze, A. and Miller, A.R. 2014. Women helping women? Evidence from private sector data on workplace hierarchies. *NBER Working Paper* No. 20761. Cambridge, MA: National Bureau of Economic Research.

Laband, D.N. and Tollison, R.D. 2003. Dry holes in economic research. *Kyklos* 56(2), 161–174.

Link, A.N., Swann, C.A. and Bozeman, B. 2008. A time allocation study of university faculty. *Economics of Education Review* 27(4), 363–374.

McDowell, J.M. and Smith, J. 1992. The effect of gender-sorting on propensity to coauthor: Implications for academic promotion. *Economic Inquiry* 30(1), 68–82.

Moers, F. 2005. Discretion and bias in performance evaluation: the impact of diversity and subjectivity. *Accounting, Organizations and Society* 30(1), 67–80.

Monks, J. and Robinson, M. 2000. Gender and racial earnings differentials in academic labor markets. *Economic Inquiry* 38(4), 662–671.

Moore, W.J., Newman, R.J. and Turnbull, G.K. 1998. Do academic salaries decline with seniority? *Journal of Labor Economics* 16(2), 352–366.

Pekkarinen, T. and Vartiainen, J. 2006. Gender differences in promotion on a Job ladder: evidence from Finnish metalworkers. *ILR Review* 59(2), 285–301.

Pema, E. and Mehay, S. 2010. The role of job assignment and human capital endowments in explaining gender differences in job performance and promotion. *Labour Economics* 17(6), 998–1009.

Pergamit, M.R. and Veum, J.R. 1999. What is a promotion?. *ILR Review* 52(4), 581–601.

Petersen, T., Snartland, V. and Meyersson Milgrom, E.M. 2007. Are female workers less productive than male workers? *Research in Social Stratification and Mobility* 25(1), 13–37.

Prendergast, C. and Topel, R. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37(2–3), 355–365.

Ransom, M.R. and Megdal, S. 1993. Sex differences in the academic labor market in the affirmative action era. *Economics of Education Review* 12(1), 21–43.

Ransom, M. and Oaxaca, R.L. 2005. Intrafirm mobility and sex differences in pay. *ILR Review* 58(2), 219–237.

Räty, T. and Bondas, M. 2008. A publication activity model for Finnish universities. *VATT Discussion Papers* No. 448. Helsinki, Finland: Government Institute for Economic Research.

Remler, D.K. and Pema, E. 2009. Why do institutions of higher education reward research while selling education? *NBER Working Paper Series* No. 14974.

Sax, L.J., Serra Hagedorn, L., Arredondo, M. and Dicrisi III, F.A. 2002. Faculty research productivity: exploring the role of gender and family-related factors. *Research in Higher Education* 43(4), 423–446.

Schneider, A. 1998. Why don't women publish as much as men. *Chronicle of Higher Education* 45(3), 14–16.

Shin, J.C. and Cummings, W.K. 2010. Multilevel analysis of academic publishing across disciplines: research preference, collaboration, and time on research. *Scientometrics* 85(2), 581–594.

Stack, S. 2004. Gender, children and research productivity. *Research in Higher Education* 45(8), 891–920.

Takahashi, A.M. and Takahashi, S. 2011. Gender salary differences in economics departments in Japan. *Economics of Education Review* 30(6), 1306–1319.

Toutkoushian, R.K. 1998. Sex matters less for younger faculty: evidence of disaggregate pay disparities from the 1988 and 1993 NCES surveys. *Economics of Education Review* 17(1), 55–71.

Toutkoushian, R.K. 1999. The status of academic women in the 1990s no longer outsiders, but not yet equals. *The Quarterly Review of Economics and Finance* 39(5), 679–698.

Toutkoushian, R.K. and Bellas, M.L. 1999. Faculty time allocations and research productivity: Gender, race and family effects. *Review of Higher Education* 22(4), 367–390.

Umbach, P.D. 2006. Gender equity in the academic labor market: an analysis of academic disciplines. *Research in Higher Education* 48(2), 169–192.

Van Herpen, M., Cools, K. and Van Praag, M. 2006. Wage structure and the incentive effects of promotions. *Kyklos* 59(3), 441–459.

Vincent-Lancrin, S. 2009. What is changing in academic research? Trends and prospects. In: *Higher Education to 2030, Volume 2: Globalisation*, OECD: Centre for Educational Research and Innovation.

Waldman, D.A. and Avolio, B.J. 1986. A meta-analysis of age differences in job performance. *Journal of Applied Psychology* 71(1), 33–38.

Ward, M. 2001. Gender and promotion in the academic profession. *Scottish Journal of Political Economy* 48(3), 283–302.

Warman, C., Woolley, F. and Worswick, C. 2010. The evolution of male-female earnings differentials in Canadian universities, 1970–2001. *Canadian Journal of Economics* 43(1), 347–372.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data* (2nd edition). The MIT Press: Cambridge, MA, US.

Xie, Y. and Shauman, K.A. 1998. Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review* 63(6), 847–870.

Appendix

Description of the variables

Variable name	Description
Monthly earnings	Monthly earnings in euros.
Female	= 1 if female, = 0 if male.
Age	Age in full years. Used as a proxy variable for potential total work experience.
Tenure	Measures the number of years of service at the university. For employees missing this information, tenure measures the length of time since the latest labor contract was negotiated; the variable will therefore underestimate actual job tenure for some employees. Furthermore, in some cases, tenure is likely to be an overestimate of actual work experience because it is measured in full years after a specified reference date and possible career breaks are not accounted for.
Education (highest degree)	Three options: master's degree (or lower), licentiate's degree, doctoral degree. Approximately 13% of the worker-year-observations lack information on education level. We imputed these missing values with the most common education level of the employees working in the same occupation. However, the reported results were essentially unchanged when individuals with missing education information were excluded from the analysis.
Occupation	Occupations: 1) doctoral student, 2) teaching assistant, 3) researcher, 4) university instructor, 5) postdoctoral researcher, 6) senior researcher, 7) senior assistant, 8) lecturer, 9) professor, 10) other occupation.
Job complexity level	11 different job complexity levels.
Number of publications	Publications are divided to three categories: 1) peer-reviewed international articles, 2) peer-reviewed national articles, 3) all other publications (e.g., book chapters, discussion papers).
Departments	27 departments.
Administrative duties	= 1 if a worker had concurrent administrative duties (i.e., earned wage bonus for administrative duties), = 0 otherwise.

TABLE A1 Task-specific and performance components of earnings in 2012

Job complexity level	Task-specific component	Performance level	Performance component (% of task-specific component)
1	1 747 €	1	0 %
2	1 922 €	2	4 %
3	2 114 €	3	10 %
4	2 403 €	4	16 %
5	2 787 €	5	22 %
6	3 254 €	6	28 %
7	3 755 €	7	34 %
8	4 543 €	8	40 %
9	5 120 €	9	46 %
10	5 796 €		
11	6 703 €		

Notes: Task-specific and performance components are based on a pay scale applied in December 2012 rounded to the nearest integer. For example, an employee working at complexity level 6 and at performance level 5 in 2012 had monthly earnings of approximately 3254 + 0.22*3254 $\approx$ 3970 euros (if he or she did not receive any additional wage bonuses).

TABLE A2 Role of gender in assignment of administrative duties

Dependent dummy variable = 1 if employee had administrative duties, = 0 otherwise		
	(1)	(2)
Female	−0.033	−0.009
	(0.010)***	(0.009)
Control variables		
Age and tenure	No	Yes
Educational level	No	Yes
Department	No	Yes
Research productivity	No	Yes
$R^2_{adjusted}$	0.01	0.12
Observations	8894	8894

Notes: Research productivity variables measure concurrent numbers of publications (international peer-reviewed articles, national peer-reviewed articles and other publications). Full results are available upon request. * Statistically significant at the .10 level; ** at the .05 level; *** at the .01 level.

TABLE A3 Marginal effects of the ordered probit models

Ordered probit models of Table 5

		Job complexity level										
		1	2	3	4	5	6	7	8	9	10	11
Column 1:	Female	0.0002**	0.0046**	0.0176***	0.0180**	−0.0081**	−0.0253**	−0.0029**	−0.0036**	−0.0005**	−0.00001*	~ 0
Column 2:	Female	~ 0	0.0003	0.0043	0.0094	0.0122	−0.0177	−0.0034	−0.0044	−0.0007	−0.00001	~ 0
Column 3:	Female	~ 0	0.0005	0.0061	0.0123	0.0150	−0.0223	−0.0043	−0.0060	−0.0011	−0.00003	~ 0

Ordered probit models of Table 6

		Individual performance level								
		1	2	3	4	5	6	7	8	9
Column 1:	Female	0.0001*	0.0006*	0.0101*	0.0152*	0.0070*	−0.0131*	−0.0145*	−0.0048*	−0.0006*
Column 2:	Female	–	0.0001	0.0017	0.0046	0.0047	−0.0024	−0.0060	−0.0024	−0.0003
Column 3:	Female	–	0.0001	0.0021	0.0056	0.0059	−0.0030	−0.0074	−0.0030	−0.0004

Notes: Marginal effects are evaluated at the means of the explanatory variables. A lack of observations is indicated by –.

CHAPTER 4: THE WAGE CURVE AND LOCAL MONOPSONY POWER *

Abstract
Using worker panel data from Finland, a country with a geographically dispersed population and relatively long distances between local labor markets, this paper examines the responsiveness of the pay level to local unemployment conditions. In particular, this study analyzes whether the pay level is more responsive to the unemployment level in less agglomerated and more remote regions, as might be expected if employers have a higher degree of local monopsony power in such regions. The results consistently suggest that the pay level is lower in localities with a higher unemployment level and hence provide strong support for the so-called wage curve hypothesis, which predicts a negative relationship between local unemployment and the pay level. Although the results provide some evidence that the magnitude of the wage curve relationship, which is estimated based on the unemployment elasticity of pay, varies across different regions of the country, the findings do not provide consistent support for the monopsony power hypothesis. In particular, once the unobserved worker heterogeneity is controlled for, the local unemployment elasticity of pay is unrelated to the degree of regional agglomeration. Further analysis based on a more direct measure of local monopsony power, namely, the number of own-industry establishments in the locality, supports a similar conclusion: the responsiveness of the pay level to local unemployment is not stronger for workers whose employers potentially have more monopsony power over them.
Keywords: wage curve, regional unemployment, monopsony power
JEL Classification: J31, J42, J60, R23

* The author would like to thank Jaakko Pehkonen, Ulrich Zierahn, Jari Vainiomäki, Jani-Petri Laamanen and the participants at the 53rd ERSA Congress (Palermo), IAB Workshop "Spatial Dimensions of the Labour Market: The Relocation of Jobs" (Nuremberg), and XXXI Summer Seminar of Finnish Economists (Jyväskylä) for helpful comments on the earlier version of this paper. This research has been financially supported by the OP-Pohjola Group Research Foundation (grant no. 201200100).

1 Introduction

Empirical literature has extensively confirmed the wage curve hypothesis proposed by Blanchflower and Oswald (1990, 1994), which suggests that wages are lower in local labor markets with higher unemployment. However, in contrast to the findings of their pioneering empirical work, the magnitude of this inverse relationship seems to vary considerably across countries (see, e.g., Nijkamp & Poot 2005). Based on their empirical findings, Blanchflower and Oswald argued that the responsiveness of wages to local unemployment is not affected by differences in the labor market institutions of countries. However, later studies have shown that local wage responsiveness is contingent on labor market institutions. Indirect evidence is based on less elastic wage curves in countries with more centralized wage bargaining systems, such as those in Nordic countries (see, e.g., Albæk et al. 2000; Nijkamp and Poot 2005), and union workers (e.g., Card 1995; Barth et al. 2002), while more direct evidence is provided by studies illustrating that the slope of the wage curve changes based on the restructuring of the wage bargaining system and other labor market re-forms (Devicienti et al. 2008; Cholezas and Kanellopoulos 2015; Daouli et al. 2017).

Furthermore, prior studies provide evidence that the slope of the wage curve may vary across different regions of the same country. For instance, Turunen (1998) found that the size and significance of the wage curve slope of young workers varies across nine major geographical areas of the US. Deller (2011) used cross-sectional regional data from the US to show that wages are negatively related to local unemployment in some parts of the country, but in others, the wage-unemployment relationship is either positive or statistically nonsignificant.

The within-country variation in wage curve estimates is somewhat puzzling: because the local labor markets of a country typically share the same wage-setting mechanisms and other labor market institutions, the regional variation in the wage curve slopes calls for an explanation that relies on other factors. One potential explanation is provided by Longhi et al. (2006), who argued that within-country variation may arise from local monopsony power caused by regional differences in geographical remoteness and in the degree of economic agglomeration. Their reasoning is as follows: the combination of fewer jobs and higher job mobility costs (including job search, commuting and migration costs) in more remote low-agglomeration regions weakens the outside job opportunities of the workers in these regions; the poorer job opportunities give employers local monopsony power over their employees, and consequently, employers achieve greater flexibility in adjusting wages based on local unemployment conditions. Consequently, the wage curve relationship is more pronounced in remote low-agglomeration regions than in regions that have a high concentration of firms and that are in close proximity to neighboring regions.

Using regional data from western Germany, Longhi et al. (2006) provided supportive empirical evidence for the monopsony power hypothesis: the slope of the wage curve is larger in less agglomerated and more isolated regions. Consistent with the monopsony power hypothesis, other studies have illustrated that the magnitude of the wage curve relationship is contingent on the degree of urbanization, where the pay level is less responsive to local unemployment level in larger cities than in smaller cities and rural regions (Baltagi et al. 2012; Baltagi and Rokicki 2014).

In this paper, we use worker-level panel data from Finland to test the monopsony power explanation of the variation in the regional wage curve. We argue that Finland provides an ideal case for analyzing local monopsony power because long distances between local labor markets create notable job mobility costs. Based on our results and conclusions made by Solon et al. (1994), we argue that the use of regional aggregate data may partly explain why Longhi et al. (2006) found a more pronounced wage curve for less agglomerated regions. Solon and others argued that using aggregate data to analyze the cyclicality of (real) wages may introduce a composition bias, leading to the underestimation of the true procyclicality of wages. This composition bias arises when aggregate statistics fail to properly account for the changing composition of the workforce over the business cycle. If the size of the composition bias varies across regions with different degrees of economic agglomeration, using regional aggregate data to analyze within-country variation in wage curve slope may lead to incorrect conclusions. As discussed in Solon et al. (1994), a solution to the composition bias problem is to estimate worker-level wage regressions while including worker fixed effects. To account for the composition bias in the analysis of regional variation in the wage curve relationship, we thus re-estimate the models estimated by Longhi et al. (2006) by using worker fixed effects specifications. Our results suggest that the greater responsiveness of the pay level to local unemployment conditions in less agglomerated regions disappears when we account for unobserved worker heterogeneity.

2 Data and empirical approach

The micro-data analyzed are based on a 7% random sample of the Finnish population drawn in 2001. The data from a sampling year were merged with data from preceding and subsequent years, and the resulting longitudinal data include information on the sampled individuals for the period from 1995 to 2006. For the purposes of the wage curve analysis, these worker-level data were combined with regional data on local unemployment rates measured at the LAU-1 level (79 sub-regions)[1]. Because the micro-data include identifiers for sub-regions only for the years 1995–2002, the final sample used for the analysis includes observations only for this period. The analysis focuses on non-

[1] The micro-data were collected by Statistics Finland. The unemployment rates were drawn from the *Employment Service Statistics* compiled by the Ministry of Economic Affairs and Employment, and they are based on the number of unemployed individuals registered as jobseekers at employment offices.

agricultural private sector employees who lived in mainland Finland, and consequently, individuals who lived in the Åland Islands (which constitute an autonomous province of Finland) and individuals who were employed in the public sector or by the agriculture, forestry or fishing industries were excluded from the final sample. Additionally, employees aged under 18 and over 68 were excluded from the sample.

To examine the regional variation in the wage curve relationship, we estimate the following earnings equation:

$$log(e_{irt}) = a + \beta\, log(u_{rt}) + \gamma\, log(u_{rt})*AM_{rt} + \lambda\, AM_{rt} + \delta\mathbf{X}_{irt} + \eta_r + \theta_t + \varepsilon_{irt} \qquad (1)$$

where i, r and t represent the individual, sub-region and year, respectively; e_{irt} is annual earnings; u_{rt} is the local unemployment rate of the LAU-1 sub-region; X_{irt} is a vector of worker characteristics, including age, age^2, work experience, work experience^2 and a set of dummy variables for gender, native/first language, marital status, children under 18 years old, education level, field of study and industry (and interaction terms of gender with marital status and children dummies); η_r is a region effect; θ_t is a year effect; and ε_{irt} is an error term. Table 1 summarizes the descriptive statistics for the worker characteristics. The regression variables are described in more detail in Table A1 in the Appendix.

Because both the earnings variable and the unemployment variable are in logarithms, coefficient β measures the local unemployment elasticity of pay. According to the wage curve relationship, the unemployment elasticity is negative; hence, we expect to observe a negative coefficient for the estimate of β. As noted by Card (1995), the relevant dimension for the estimation of the unemployment coefficient β is the product of the number of regions and the number of observation years. Consequently, the estimation of β is effectively based on 632 observations (= 79 regions * 8 years).

Variable AM_{rt} captures the degree of economic agglomeration of sub-region r in year t and is measured in three different ways. The first two measures are adopted from Longhi et al. (2006) and are calculated as follows[2]:

$$A_{rt} = \sum_i 10^{-6} * (E_{it}*w_{ri})$$

$$T_{rt} = \sum_i T_{irt} = \sum_i 10^{-6} * (E_{it}*E_{rt})^{0.5} * w_{ri}$$

where E_{rt} is the total number of employed in sub-region r in year t; E_{it} is the total number of employed in region i neighboring sub-region r; and weights w_{ri} are elements of the R x R spatial weight matrix, where R is the number of sub-regions (the rows and columns of the matrix have the same ordering of sub-regions). The elements of the spatial weight matrix are calculated based on the inverses of Euclidean distances between sub-region r and its neighboring sub-

2 For a more detailed description of these measures, see Longhi et al. (2006), pages 716–720.

regions i, and diagonal elements (w_{rr}) and elements corresponding to non-contiguous sub-regions (i.e., regions that do not border one another) are set to zero[3]. The weight matrix is "row standardized" by dividing each element of the matrix by the row sum of the elements. Consequently, each row of the final spatial weight matrix sums to zero; that is, for each sub-region r, $\Sigma_i w_{ri} = 1$. Measure A_{rt} is hence a weighted average of the number of employed individuals in sub-regions surrounding sub-region r. Assuming that local monopsony power yields a more elastic wage curve for more remote and less agglomerated regions, we would expect to observe a negative value for parameter γ. Table 1 reports the summary statistics of the agglomeration measures A_{rt} and T_{rt}.

TABLE 1 Descriptive statistics (1995–2002)

Variable	Mean	Std. Dev.	Min	Max
Worker characteristics				
Annual earnings e_{irt} (euros)	24398	8989	1900	69000
Age	39.9	10.2	18	68
Work experience	10.2	3.5	1	16
Female dummy	0.422			
Marital dummy	0.540			
Children dummy	0.427			
Education level				
Primary/lower secondary	0.253			
Upper secondary	0.446			
Lowest level tertiary	0.183			
Lower-degree level tertiary	0.063			
Higher-degree level tertiary	0.053			
Doctorate or equivalent	0.002			
Regional variables				
Regional unemployment rate u_{rt}	16.76	4.75	6.38	30.22
Regional agglomeration measure A_{rt}	0.027	0.029	0.008	0.206
Regional agglomeration measure T_{rt}	0.019	0.016	0.006	0.139

Notes: The marital dummy equals one if a worker was married and zero otherwise. The children dummy equals one if a worker had children under 18 years old and zero otherwise.

As a third measure of agglomeration variable AM_{rt} in earnings equation (1), we use a set of dummy variables that divide sub-regions into geographical areas with different levels of agglomeration. Based on regional characteristics related

[3] Because the sub-regions comprise two or more municipalities and have more than one local center, the calculation of the elements of the spatial weight matrix is based on the Euclidean distances between the administrative centers of the most populated municipalities within each sub-region. The data on Euclidean distances were extracted from Google Maps.

to the degree of economic agglomeration reported in Table 2, three dummy variables are generated for the geographical areas of Finland: "South" (high-agglomeration region), "Central" (medium-agglomeration region) and "North" (low-agglomeration region)[4]. In this model specification, the southern sub-regions are used as a reference group, and we are particularly interested of the coefficient estimate on the interaction term between the local unemployment variable and the dummy variable for northern sub-regions. A north-south division provides an ideal setting for a test of the monopsony power hypothesis as an explanation for the within-country variation in wage curve estimates: the population in Finland is heavily concentrated in the southern parts of the country, whereas northern Finland, characterized by a lack of economic agglomerations and long distances between local labor markets, is among the least populated geographical areas in Europe. Hence, based on the monopsony power hypothesis, we would expect to observe a more pronounced wage curve for the northern sub-regions than for the southern sub-regions.

TABLE 2 Region characteristics in 2002

Variable	Region "South"	"Central"	"North"
Population share (%)	46.3	37.0	16.7
Population/km² of land area	32.0	14.6	4.3
Land area (%)	18.4	32.1	49.5
Degree of urbanization (%)[a]	84.7	74.5	67.7
Mean of the Euclidean distance between the sub-regional centers (km)	57.6	67.2	105.8
Number of municipalities			
< 10 000 inhabitants	141	132	56
10000–49999 inhabitants	42	30	15
50000–99999 inhabitants	3	5	0
≥ 100000 inhabitants	5	0	1

Notes: [a] Mean of values of included NUTS-3 regions. In 2002, the total population of Finland was 5,180,038, and the total land area was 302,946 km² (excluding the population and the land area of Åland Islands). Source: Author's own calculations based on data by Statistics Finland.

4 "North" consists of the sub-regions located in the three northernmost NUTS-3 regions (Kainuu, Northern Ostrobothnia and Lapland). "South" comprises the sub-regions located in the eight contiguous NUTS-3 regions in south-western Finland (Uusimaa, Itä-Uusimaa, Varsinais-Suomi, Satakunta, Kanta-Häme, Pirkanmaa, Päijät-Häme and Kymenlaakso), constituting the most densely populated area in the country. "Central" comprises the sub-regions located in the eight NUTS-3 regions of the central Finland (South Karelia, Etelä-Savo, Pohjois-Savo, North Karelia, Central Finland, South Ostrobothnia, Ostrobothnia and North Ostrobothnia).

The dependent variable of the earnings equation (1) is a logarithm of worker's annual earnings. The wage curve relationship essentially describes the relationship between the local unemployment level and the *wage level*, and hence, the preferable dependent variable would be workers' hourly wage. Unfortunately, the micro-data used for the analysis do not include information on hourly wages or working hours. When annual earnings are used as a dependent variable for the regression equation, the coefficient on the unemployment variable, β, may yield inflated estimates of the true wage-unemployment relationship, as β reflects the variation in both hourly wages and annual working hours with respect to local unemployment level. To account for the effects of varying working hours on the wage curve estimates, the final sample used for the analysis excludes 1) employees who worked less than twelve months a year and 2) 5% of the observations from both tails of the annual earnings distribution in each year in each sub-region. The final sample employed in the estimation of different specifications of earnings equation (1) consists of 429,414 observations for 92,839 workers.

Earnings equation (1) is estimated by using an ordinary least squares estimator. Additionally, to account for the pay-level effects of unobserved time-invariant worker characteristics, the earnings equation is also estimated by using the fixed effects estimator. In the fixed effects specifications, a worker-specific effect τ_i is included as an additional regressor. Earnings equations that include the worker effects are our preferred specifications, as they account for the potential composition bias that may arise from the compositional changes in the workforce over the business cycle (Solon et al. 1994). To account for the possibility that the different level of aggregation of the dependent and the independent variable, where earnings are measured at the worker level and unemployment is measured at the regional level, may cause error terms ε_{irt} to be correlated across employees working in the same sub-region (Moulton 1986, 1990), we cluster standard errors at the regional level (see Cameron and Miller 2011, 2015).

When we estimate the wage curve relationship by using the micro-data, we assume that the logarithm of local unemployment rate is an exogenous variable. While some studies use regional data to show that simultaneity bias, which is caused by the simultaneous determination of local wage and unemployment levels, may have a substantial effect on wage curve estimates (Baltagi and Blien 1998, Baltagi et al. 2000, Longhi et al. 2006), this may be less of a problem when the wage curve is estimated by using micro-data: local unemployment rates are presumably affected by aggregate wages, not by individual wages (Nijkamp and Poot 2005).

3 Results

3.1 Wage curve estimates from regionally aggregated data

To ensure the comparability of our results with those of Longhi et al. (2006), who employed regional data instead of micro-data, we begin our analysis by estimating regional-level wage curve regressions that closely correspond to their estimates. For this purpose, we first aggregate the micro-data by sub-region (i.e., we calculate the mean values of the worker-level variables for each sub-region in each year) and then use these data to estimate earnings regressions that include the same explanatory variables of interest that Longhi and others used in their regression models. These earnings regressions also include the following control variables for sub-regions: the mean age of workers, mean work experience of workers, share of female workers, share of highly educated workers (tertiary or doctorate degree or equivalent), share of workers with a degree in science and share of manufacturing workers. The earnings regressions are first estimated by using the standard OLS estimator, with region fixed effects included. Next, following Longhi and others, the two-stage least squares (2SLS) estimator is used to account for the possibility that local pay and unemployment levels may have been simultaneously determined. In the 2SLS estimations, the logarithm of local unemployment rate is instrumented with its one-year lagged value.

The results of the earnings regressions estimated using the regionally aggregated data are presented in Table 3. For brevity, only the coefficient estimates of the explanatory variables of interest are reported; the full results are available upon request. The estimates in the first column confirm the existence of the wage curve relationship in Finland, indicating that the pay level is negatively related to the local level of unemployment[5]. The OLS regression results provide a smaller unemployment coefficient (−0.065) than the two-stage least squares regression results (−0.099). The latter coefficient estimate is identical to the uniform estimate of −0.1 reported by Blanchflower and Oswald (1994) for twelve countries. Furthermore, it closely corresponds to the wage curve estimate found by Maczulskij (2013) for Finnish private sector workers based on micro data.

The specifications in the second column incorporate a variable named 'Neighboring unemployment', which measures the unemployment level in neighboring sub-regions. The value of this variable for a particular sub-region is a logarithm of a weighted sum of the unemployment rates in neighboring sub-regions, where the weights are the elements of the spatial weight matrix. The coefficient estimate on this variable is negative in both specifications, but it is

---

5 Longhi et al. (2006) also included a squared term of the logarithmic unemployment rate as an explanatory variable and observed a significant positive coefficient estimate for this variable. Our preliminary estimations consistently provided statistically nonsignificant coefficient estimates for the squared term, and it was hence excluded from the final specifications.

statistically significant only in the OLS regression (at the 10% level). Hence, the results of the 2SLS regression suggest that unemployment conditions in neighboring sub-regions play no role in determining the pay level. This finding contradicts the results of Longhi et al. (2006), who found that the neighboring unemployment level is a statistically significant determinant of the pay level. The contradiction between our findings and theirs may be attributable to differences in the regional disaggregation of data. While they used 327 regions of Western Germany to define local labor markets, we use a regional classification that disaggregates Finland into 79 sub-regions[6]. The geographical disaggregation used by Longhi and others may yield regions that are effectively too small to characterize functional local labor markets. In such a case, the pay level of a particular region is likely to be affected by not only the prevailing unemployment level in that region but also the unemployment conditions in neighboring regions (as these regions are within commuting distance for workers and hence effectively constitute a part of their local labor market). The regions used in our analysis, on the other hand, are typically considered reasonable approximations of local labor markets. Unfortunately, our data lack information on identifiers of more disaggregated regions, and we are hence unable to test the effects of different regional disaggregations on our wage curve estimates.

The specifications in columns (3)–(5) assess the regional variation in the wage curve relationship by including interaction terms for the local unemployment variable and different agglomeration measures specified in the previous section[7]. The results in columns (3) and (4) confirm the findings of Longhi et al. (2006): the coefficient estimates on the interaction terms of agglomeration variables A_{rt} and T_{rt} are statistically significant and negative, indicating that the wage curve relationship is more pronounced in less agglomerated regions. Furthermore, the estimates in column (5) indicate that the slope of the wage curve is substantially larger in less agglomerated northern sub-regions than in highly agglomerated southern sub-regions, and this difference is statistically significant. Overall, the wage curve slopes from the earnings regressions estimated by using the regionally aggregated data provide strong support for the monopsony power hypothesis, according to which the responsiveness of wages to the local unemployment level is stronger in regions with a low degree of economic agglomeration.

6 The total geographical area of western Germany is about 74% of that of Finland.

7 The coefficient estimate on the neighboring unemployment variable is statistically nonsignificant in these specifications, and the variable was therefore excluded. Furthermore, and in contrast to findings of Longhi et al. (2006), our estimations also provide a statistically nonsignificant coefficient estimate on the interaction term for the local unemployment variable and neighboring unemployment variable.

TABLE 3 Wage curve estimates (regional data)

Dependent variable: log(annual earnings)					
			Agglomeration measures		
	Baseline model 1	Baseline model 2	Agglomeration measure A_{rt}	Agglomeration measure T_{rt}	Area dummies
OLS					
$\log(u_{rt})$	−0.065***	−0.046***	−0.079***	−0.087***	−0.075***
Neighboring unemployment$_{rt}$		−0.044*			
$\log(u_{rt})$*A_{rt}			0.724**		
$\log(u_{rt})$*T_{rt}				3.422***	
A_{rt}			0.603**		
T_{rt}				5.537***	
$\log(u_{rt})$*Central					0.034***
$\log(u_{rt})$*North					−0.045**
IV (2SLS)					
$\log(u_{rt})$	−0.099***	−0.076***	−0.150***	−0.133***	−0.108***
Neighboring unemployment$_{rt}$		−0.052			
$\log(u_{rt})$*A_{rt}			1.395***		
$\log(u_{rt})$*T_{rt}				3.872***	
A_{rt}			0.278		
T_{rt}				4.835***	
$\log(u_{rt})$*Central					0.035***
$\log(u_{rt})$*North					−0.052***
Region characteristics	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Observations (OLS)	632	632	632	632	632
Observations (2SLS)	553	553	553	553	553
$R^2_{adjusted}$ (OLS)	0.98	0.98	0.98	0.98	0.98
$R^2_{adjusted}$ (2SLS)	0.98	0.98	0.98	0.98	0.98

Notes: All models include the following control variables for the sub-regions: the mean age of workers, mean work experience of workers, share of female workers, share of highly educated workers (tertiary or doctorate degree or equivalent), share of workers with a degree in science and share of manufacturing workers. In the 2SLS specifications, the logarithmic unemployment variable is instrumented with its one-year lagged value. Robust standard errors are presented in parentheses. Significant at the * 10% level; ** 5% level; *** 1% level.

3.2 Wage curve estimates from the micro-data

Next, to make more efficient use of the micro-data, we estimate the alternative specifications of worker-level earnings equation (1). The main results of these estimations are reported in Table 4.[8] The specifications reported in the table are parallel to those estimated in Table 3 using regionally aggregated data, and hence, they allow a direct comparison of the regression coefficients of interest. Again, the coefficient estimates on the unemployment variable confirm the existence of the negatively sloping wage curve relationship. However, the estimates demonstrate that the slope of the wage curve is highly sensitive to inclusion of worker effects: when unobserved worker heterogeneity is controlled for, the unemployment coefficient increases from −0.019 to −0.089. Hence, the findings highlight the importance of controlling for worker fixed effects when estimating the local pay-unemployment relationship; otherwise, the true unemployment elasticity of pay may be obscured. The latter estimate suggests that an increase in the local unemployment rate by 100% reduces the pay level by approximately 9%. The results in column (2) once again confirm that the pay level is not related to unemployment conditions in neighboring sub-regions.

The wage curve estimates in columns (3)–(5) challenge the previous findings of Longhi et al. (2006) for western Germany: after including worker fixed effects, the coefficient estimates on the interaction terms for the local unemployment variable and alternative agglomeration measures are no longer statistically significant. In other words, the steeper slope of the wage curve of the less agglomerated sub-regions disappears when unobserved worker heterogeneity is controlled for. Hence, the results do not support the hypothesis that monopsonistic features of more remote and less agglomerated regions generate a more pronounced wage curve relationship for these regions.

The use of regional aggregate data may partly explain why Longhi et al. (2006) found a larger wage curve slope for regions with a lower degree of economic agglomeration. As noted by Solon et al. (1994), using aggregate data to analyze the responsiveness of wages to unemployment conditions may introduce a composition bias, leading to the underestimation of the true procyclicality of wages. Composition bias arises when aggregate statistics fail to properly control for the changing composition of the workforce over the business cycle. Provided that composition bias is more pronounced in highly agglomerated regions, one may observe a smaller wage curve slope for these regions if the composition bias is not accounted for. As discussed in Solon et al. (1994), a solution to the composition bias problem is to estimate a micro-level wage equation while including worker fixed effects. Our results suggest that once worker fixed effects are included, the wage curve slopes are similar across regions with different degrees of economic agglomeration. Adjusting for worker composition effects may also explain why Baltagi et al. (2012) found only slightly larger

---

8 To save space, only the coefficients of interest are reported. Table A2 in the appendix reports coefficient estimates on control variables of the specifications estimated in the first column.

wage curve estimates for Western German regions in rural areas than for regions with large cities.

Although the worker fixed effects specification in column (5) yields a similar slope estimate of the wage curve for the southern and the northern sub-regions of Finland (approximately −0.09), the estimates suggest that the wage curve slope is smaller for the sub-regions located in the central parts of the country (−0.06). Hence, the results imply that the magnitude of the wage curve relationship is not uniform across different regions. In contrast to findings of Deller (2011) for the local labor markets of the US, our estimates do not provide evidence for a positively sloping wage-unemployment relationship in any of the geographical areas; rather, the coefficient estimates indicate that the pay level is inversely related to the local unemployment level in each geographical area ("North", "Central" and "South"), providing strong evidence in favor of the wage curve relationship. Based on the discussion in Card (1995), Deller's cross-sectional results may be partly driven by the lack of region fixed effects: without controls for regions, the estimated coefficient on the unemployment variable may reflect the potentially positive relationship between the "permanent" regional unemployment level and "permanent" wages, whereas region fixed effects must be included to produce the negative relationship between actual local unemployment and actual wages.

TABLE 4 Wage curve estimates (micro data)

	Dependent variable: log(annual earnings)				
	Baseline model 1	Baseline model 2	Agglomeration measure A_{rt}	Agglomeration measure T_{rt}	Area dummies
OLS					
$\log(u_{rt})$	−0.019**	−0.048**	−0.019*	−0.053***	−0.025***
Neighboring unemployment$_{rt}$		0.056			
$\log(u_{rt})$*A_{rt}			0.044		
$\log(u_{rt})$*T_{rt}				0.985**	
A_{rt}			0.132		
T_{rt}				1.079**	
$\log(u_{rt})$*Central					0.015
$\log(u_{rt})$*North					−0.075***
Worker FE					
$\log(u_{rt})$	−0.089***	−0.088***	−0.101***	−0.035***	−0.085***
Neighboring unemployment$_{rt}$		−0.003			
$\log(u_{rt})$*A_{rt}			0.183		
$\log(u_{rt})$*T_{rt}				0.383	
A_{rt}			−0.470***		
T_{rt}				2.051***	
$\log(u_{rt})$*Central					0.025***
$\log(u_{rt})$*North					−0.007
Worker characteristics	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
OLS					
$R^2_{adjusted}$	0.44	0.44	0.44	0.44	0.44
Worker FE					
R^2 (within)	0.40	0.40	0.40	0.40	0.40
R^2 (between)	0.31	0.31	0.31	0.30	0.31
R^2 (overall)	0.31	0.31	0.31	0.30	0.31
F-test (worker FE)	14.69	14.68	14.68	14.68	14.68
(*p*-value)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)

Notes: The number of worker-year-observations is 429,414 (on 92,839 individuals) and the number of region-year-observations is 632 (= 79 regions*8 years). Worker characteristics include the following variables: gender, age, age^2, work experience, work experience^2, marital status (and its interaction with gender), children dummy (and its interaction with gender), native/first language dummies, education level dummies, field of education dummies, industry dummies and year dummies (gender and language dummies are omitted in the worker fixed effects specifications). Robust Hausman tests reject random effects in favor of fixed effects for all model specifications. The *F*-statistics test the joint statistical significance of the worker fixed effects. Significant at the * 10% level; ** 5% level; *** 1% level (the significance levels are based on cluster-robust standard errors, clustered at the regional level).

3.3 Additional findings

In the previous analysis, the working hypothesis is that because of the existence of fewer potential employers, employees in the less agglomerated regions have poorer outside job opportunities, which gives local employers monopsony power that they can take advantage of when adjusting wages to local unemployment conditions. However, the degree of regional agglomeration may be a poor measure of local job opportunities, at least for some employees. For example, a low-agglomeration region may have a high concentration of firms operating in a certain industry, improving the local outside job opportunities of the employees working in these firms (or, conversely, high-agglomeration regions may have only few firms operating in certain industries, leaving less outside job opportunities for the employees working in these firms). Hence, to examine the role of local monopsony power as a determinant of the wage curve slope in more detail, a more direct measure of local job opportunities is needed.

One such measure is the number of establishments operating in the same industry and region as workers' current employer (hereafter referred to as "local own-industry establishments")[9]. In the following analysis, we employ the number of local own-industry establishments to provide a more detailed examination of the effect of local monopsony power on the slope of the wage curve relationship. The number of own-industry establishments is likely to be a more relevant indicator of local monopsony power for some employees than others. In particular, employees in some occupations (e.g., employees in maintenance, clerical and low-level administrative occupations) may more easily change working industries than employees in occupations that require more industry-specific education and training[10]. Furthermore, as job opportunities essentially depend on the amount of open vacancies, employees' outside job opportunities may be limited even in localities with a vast number of own-industry establishments, if only few of the establishments are concurrently offering vacancies (Manning 2003). However, the number of open vacancies is likely to be positively correlated with the number of establishments.

Table 5 presents estimates from the earnings regressions that include the logarithm of the number of own-industry establishments and its interaction term with the unemployment variable as explanatory variables. All estimated specifications include dummy variables for two-digit industries; otherwise, establishment count variables might also capture other industry pay effects. The specifications in columns (1) and (2) are based on the establishment count data for sub-regions, which contain the number of establishments for 25 manufacturing industries. Compared with the regional classification used in the previous

---

9 Previously, Muehlemann et al. (2013) used the number of local own-industry establishments to proxy for the degree of local monopsony power of employers.

10 Related studies have analyzed the role of employer monopsony power in the pay determination of occupations that require specific education/qualifications, such as nurses (Hirsch and Schumacher 1995, 2005), teachers (Merrifield 1999) and university faculty (Ransom 1993), whose outside job opportunities are typically limited by the small number of potential local employers (hospitals, schools and universities, respectively).

analysis, these data are based on a revised LAU-1 classification that disaggregates Finland into 74 sub-regions (instead of 79 sub-regions)[11]. The specifications in columns (3) and (4) are estimated by using unemployment statistics and establishment count data for NUTS-3 regions (19 regions). The advantage of using the NUTS-3-level data is that they include the number of local establishments for a wider range of industries; the data used for the analysis include establishment counts for 52 industries. Furthermore, using NUTS-3 regions instead of LAU-1 sub-regions allows us to employ the full micro-data, that is, data for the period from 1995 to 2006.

Most of the specifications reported in Table 5 yield a positive and statistically significant coefficient estimate on the establishment count variable, suggesting that the pay level increases with the number of local own-industry establishments. This finding is consistent with the hypothesis that a smaller number of potential employers in the locality reduces the outside job opportunities of workers, giving local employers monopsony power over their employees and consequently allowing employers to pay lower wages. The existence of economies of agglomeration provides another potential explanation for the positive relationship between the pay level and establishment count variable: the regional concentration of firms (industries) may create productivity gains and consequently higher wages (e.g., Ellison et al. 2010; Glaeser 2010)[12]. Furthermore, the increase in the number of same-industry firms in the locality may increase competition for workers, resulting in higher wages for the employees working in these firms.

The interaction term for the unemployment variable and the establishment count variable is generally statistically nonsignificant and close to zero, hence contradicting the hypothesis that the magnitude of the wage curve relationship depends on the degree of employer monopsony power. An exception is the worker fixed effects specification in column (4), which provides a significant negative coefficient on the interaction term, suggesting that the slope of the wage curve increases with the number of own-industry establishments.

---

11 The reduction in the number of sub-regions by five is attributable to mergers of the contingent sub-regions. To estimate the specifications in columns (1) and (2), we had to modify the sub-region identifiers of the micro-data in order for them to be compatible with the regional classification of the establishment count data. This modification resulted in inaccurate sub-region identifier for some individuals. However, the number of individuals with misspecified identifiers is small, and the measurement error is hence expected to have a negligible effect on the estimation results. We tested the robustness of the results by re-estimating the models in columns (1) and (2) by using a restricted sample that excluded all observations for the sub-regions that potentially included individuals with misspecified sub-region identifiers. These estimations yielded coefficient estimates very similar to those reported in Table 5.

12 The concentration-related productivity gains arise from the close geographical proximity of firms, which, for example, improves the supply chains of the firms and increases the interaction of firms and flows of workers, technology and information between firms.

TABLE 5 The wage curve and local number of establishments

Dependent variable: log(annual earnings)				
Regional level (# of regions)	*LAU-1* (74 regions)		*NUTS-3* (19 regions)	
Estimation period	1995–2002		1995–2006	
Number of 2-digit industries	25		52	
OLS				
$\log(u_{rt})$	−0.023	0.004	0.008	0.030
	(0.014)	(0.058)	(0.016)	(0.014)**
log(own-industry establishments$_{rt}$)	0.022	0.033	0.036	0.043
	(0.007)***	(0.022)	(0.008)***	(0.005)***
$\log(u_{rt})$*log(own-industry establishments$_{rt}$)		−0.004		−0.003
		(0.007)		(0.002)
Worker fixed effects				
$\log(u_{rt})$	−0.046	−0.045	−0.058	0.041
	(0.007)***	(0.023)**	(0.010)***	(0.012)***
log(own-industry establishments$_{rt}$)	−0.007	−0.007	0.012	0.038
	(0.004)**	(0.009)	(0.003)***	(0.004)***
$\log(u_{rt})$*log(own-industry establishments$_{rt}$)		−0.0002		−0.012
		(0.003)		(0.001)***
Industry dummies (2-digit)	Yes	Yes	Yes	Yes
Worker characteristics	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Worker-year observations	163,146	163,146	679,268	679,268
OLS				
$R^2_{adjusted}$	0.48	0.48	0.50	0.50
Worker FE				
R^2 (within)	0.40	0.40	0.48	0.48
R^2 (between)	0.29	0.29	0.36	0.36
R^2 (overall)	0.29	0.29	0.37	0.36
F-test (worker FE)	13.42	13.42	14.28	14.30
(*p*-value)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)

Notes: Cluster-robust standard errors in parentheses, clustered at the regional level. The mean value (standard deviation) of the non-logarithmic establishment count variable is 25.2 (57.8) in the LAU-1 level establishment data and 225.2 (671.9) in the NUTS-3 level establishment data. Worker characteristics include the following variables: gender, age, age^2, work experience, work experience^2, marital status (and its interaction with gender), children dummy (and its interaction with gender), native/first language dummies, education level dummies, field of education dummies and year dummies (gender and language dummies are omitted in the worker fixed effects specifications). Robust Hausman tests reject random effects in favor of fixed effects for all model specifications. The *F*-statistics test the joint statistical significance of the worker fixed effects. Significant at the * 10% level; ** 5% level; *** 1% level.

4 Conclusions

This paper employs longitudinal micro-data on private sector workers and regional data on 79 local labor markets from Finland to examine the within-country variation in the local unemployment elasticity of pay. The results provide strong support for the existence of the so-called wage curve relationship, which states that the wage level decreases with the regional unemployment rate. Furthermore, the results indicate that conditional on the local unemployment rate, the unemployment conditions in neighboring regions do not play a role in determining the pay level.

The results provide some evidence that the slope of the wage curve varies across different geographical areas of a country. Moreover, the findings indicate that once worker fixed effects are included to control for the composition bias resulting from the changing composition of the workforce (Solon et al. 1994), wage curve slopes are similar across regions with different degrees of economic agglomeration. Hence, the findings contradict the monopsony power hypothesis proposed by Longhi et al. (2006), which predicts that the magnitude of the wage curve relationship is stronger in less agglomerated regions because of the higher monopsony power of employers in these regions, and imply that the failure to control for the unobserved worker heterogeneity (composition bias) may explain why Longhi and others found a more pronounced wage curve relationship for the low-agglomeration regions than for the high-agglomeration regions of Western Germany. Further analysis based on a more direct measure of local monopsony power, namely, the number of own-industry establishments in the locality, yields a similar conclusion: the pay responsiveness to local unemployment conditions is not stronger for employees whose employers potentially have more monopsony power over them.

References

Albæk, K., Asplund, R., Barth, E., Blomskog, S., Guômundsson, B.R., Karlsson, V. and Strøjer Madsen, E. 2000. Dimensions of the wage-unemployment relationship in the Nordic countries: Wage flexibility without wage curves. In: Polachek, S. and K. Tatsiramos (ed.) *Research in Labor Economics*, Volume 19, pp. 345–381.

Baltagi, B.H. and Blien, U. 1998. The German wage curve: evidence from the IAB employment sample. *Economics Letters* 61(2), 135–142.

Baltagi, B.H., Blien, U. and Wolf, K. 2000. The East German wage curve 1993–1998. *Economics Letters* 69(1), 25–31.

Baltagi, B.H., Blien, U. and Wolf, K. 2012. A dynamic spatial panel data approach to the German wage curve. *Economic Modelling* 29(1), 12–21.

Baltagi, B. H. and Rokicki, B. 2014. The spatial Polish wage curve with gender effects: Evidence from the Polish Labor Survey. *Regional Science and Urban Economics* 49(1), 36–47.

Barth, E., Bratsberg, B., Naylor, R.A. and Raaum, O. 2002. Explaining variations in wage curves: Theory and evidence. University of Oslo, Department of Economics, *Memorandum* No. 03/2002.

Blanchflower, D.G. and Oswald, A.J. 1990. The wage curve. *Scandinavian Journal of Economics* 92(2), 215–237.

Blanchflower, D.G. and Oswald, A.J. 1994. *The Wage Curve*. The MIT Press: Cambridge, MA, US.

Cameron, A.C. and Miller, D.L. 2011. Robust inference with clustered data. In: Ullah, A. and Giles, D.E.A. (eds.) *Handbook of Empirical Economics and Finance*, pp. 1–28. Boca Raton, FL, US: Chapman & Hall/CRC.

Cameron, A.C. and Miller, D.L. 2015. A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317–372.

Card, D. 1995. The wage curve: A review. *Journal of Economic Literature* 33(2), 785–799.

Cholezas, I. and Kanellopoulos, N.C. 2015. Labour market reforms in Greece and the wage curve. *Economics Letters* 136(1), 19–21.

Daouli, J., Demoussis, M., Giannakopoulos, N. and Laliotis, I. 2017. The wage curve before and during the Greek economic crisis. *Empirical Economics* 52(1), 59–77.

Deller, S. 2011. Spatial heterogeneity in the wage curve. *Economics Letters* 113(3), 231–233.

Devicienti, F., Maida, A. and Pacelli, L. 2008. The resurrection of the Italian wage curve. *Economics Letters* 98(3), 335–341.

Ellison, G., Glaeser, E.L. and Kerr, W.R. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 100(3), 1195–1213.

Glaeser, E.L. 2010. *Agglomeration Economics*. National Bureau of Economic Research Conference Report. The University of Chicago Press: Chicago, IL, US.

Hirsch, B.T. and Schumacher, E.J. 1995. Monopsony power and relative wages in the labor market for nurses. *Journal of Health Economics* 14(4), 443–476.

Hirsch, B.T. and Schumacher, E.J. 2005. Classic or new monopsony? Searching for evidence in nursing labor markets. *Journal of Health Economics* 24(5), 969–989.

Longhi, S., Nijkamp, P. and Poot, J. 2006. Spatial heterogeneity and the wage curve revisited. *Journal of Regional Science* 46(4), 707–731.

Manning, A. 2003. The real thin theory: monopsony in modern labour markets. *Labour Economics* 10(2), 105–131.

Maczulskij, T. 2013. Public–private sector wage differentials and the business cycle. *Economic Systems* 37(2), 284–301.

Merrifield, J. 1999. Monopsony power in the market for teachers: Why teachers should support market-based education reform. *Journal of Labor Research* 20(3), 377–391.

Moulton, B.R. 1986. Random group effects and the precision of regression estimates. *Journal of Econometrics* 32(3), 385–397.

Moulton, B.R. 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics* 72(2), 334–338.

Muehlemann, S., Ryan, P. and Wolter, S.C. 2013. Monopsony power, pay structure, and training. *ILR Review* 66(5), 1097–1114.

Nijkamp, P. and Poot, J. 2005. The last word on the wage curve?. *Journal of Economic Surveys* 19(3), 421–450.

Ransom, M.R. 1993. Seniority and monopsony in the academic labor market. *American Economic Review* 83(1), 221–233.

Solon, G., Barsky, R. and Parker, J.A. 1994. Measuring the cyclicality of real wages: How important is composition bias?. *The Quarterly Journal of Economics* 109(1), 1–25.

Turunen, J. 1998. Disaggregated wage curves in the United States: Evidence from panel data of young workers. *Applied Economics* 30(12), 1665–1677.

Appendix

TABLE A1 Description of the variables

Variable	Description
Annual earnings	Annual earnings in euros.
Regional unemployment rate	Regional unemployment rate, computed at LAU-1 level (79 sub-regions). The sources of the unemployment data: Labour Force Survey (LFS), Statistics Finland.
Gender	Female dummy: = 1 if female, = 0 if male
Age	Age in full years
Experience	Work experience in years (since 1987). The values are calculated as follows: (The number of working months since January 1987) / 12
Marital status	Marriage dummy: = 1 if married, = 0 otherwise
Native language	Dummy variables for native/first language: 1) Finnish, 2) Swedish or 3) other
Children	Children dummy: = 1 if the worker had children under 18 years, = 0 otherwise
Level of education (based on ISCED 1997)	1) Primary education or lower secondary education (or unknown), 2) Upper secondary level education, 3) Lowest level tertiary education, 4) Lower-degree level tertiary education, 5) Higher-degree level tertiary education, 6) Doctorate or equivalent level tertiary
Field of study (based on ISCED 1997 classification)	1) General programmes (or not known or unspecified), 2) Education, 3) Humanities and arts, 4) Social sciences, business and law, 5) Science, 6) Engineering, manufacturing and construction, 7) Agriculture, 8) Health and welfare, 9) Services
Industry (based on NACE classification)	1) Agriculture, forestry and fishing (excluded from the analysis), 2) Mining and quarrying, 3) Manufacturing, 4) Electricity, gas and water supply, 5) Construction, 6) Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods, 7) Hotels and restaurants, 8) Transport, storage and communication, 9) Financial intermediation, 10) Real estate, renting and business activities, 11) Public administration and defence; compulsory social security, 12) Education, 13) Health and social work, 14) Other community, social and personal service activities, 15) Private households employing domestic staff and undifferentiated production activities of households for own use; extra-territorial organizations and bodies; industry unknown

Notes: Source of education and industry classifications: Statistics Finland.

TABLE A2 Earnings equations

Dependent variable: log(annual earnings)				
	OLS		Worker fixed effects	
	Coefficient	Robust SE	Coefficient	Robust SE
log(u_{rt})	−0.019	(0.008)**	−0.089	(0.012)***
Female	−0.168	(0.010)***		
Age	0.016	(0.004)***	0.037	(0.004)***
Age^2	−0.0002	(0.00005)***	−0.0004	(0.00006)***
Work experience	0.036	(0.001)***	0.085	(0.004)***
Work experience^2	−0.0005	(0.0001)***	−0.002	(0.00008)***
Married	0.043	(0.005)***	0.012	(0.002)***
Married*Female	−0.054	(0.006)***	−0.047	(0.004)***
Children dummy	0.0003	(0.003)	−0.010	(0.002)***
Children dummy*Female	−0.039	(0.008)***	−0.079	(0.012)***
Education level				
Primary/lower secondary	(omitted)		(omitted)	
Upper secondary	0.084	(0.013)***	−0.009	(0.013)
Lowest level tertiary	0.210	(0.017)***	0.109	(0.030)***
Lower-degree level tertiary	0.336	(0.025)***	0.221	(0.017)***
Higher-degree level tertiary	0.466	(0.024)***	0.350	(0.024)***
Doctorate or equivalent	0.489	(0.016)***	0.452	(0.028)***
Language dummies				
Finnish	(omitted)			
Swedish	−0.014	(0.019)		
Other/unknown	0.006	(0.008)		
Other controls				
Field of study	Yes		Yes	
Industry	Yes		Yes	
Year	Yes		Yes	
Region	Yes		Yes	
Worker-year observations	429,414		429,414	
Workers	92,839		92,839	
$R^2_{adjusted}$	0.44			
R^2 (within)			0.40	
R^2 (between)			0.31	
R^2 (overall)			0.31	

Notes: These are the same model specifications estimated in column (1) of Table 2. Robust SE = cluster-robust standard errors, clustered at the regional level. Children dummy equals one if a worker had children under 18 years old and zero otherwise. Significant at the * 10% level; ** 5% level; *** 1% level.

CHAPTER 5: ALTERNATIVE UNEMPLOYMENT MEASURES AS PREDICTORS OF REGIONAL PAY DIFFERENCES *

Abstract

Using worker-level panel data from Finland, this paper examines the implications of using alternative measures of local unemployment to estimate the regional unemployment elasticity of pay. The empirical analysis consistently yields negative elasticity estimates, suggesting that local unemployment level is negatively related to pay level. However, the results illustrate that a statistically and economically significant inverse pay-unemployment relationship (commonly referred to as the wage curve) is typically only detected when worker fixed effects are controlled for. The estimate of the local unemployment elasticity of pay varies considerably depending on the unemployment measure used in the analysis, but is not sensitive to the level of geographical disaggregation at which unemployment is measured. Modifying standard unemployment rates to exclude long-term unemployed and to include participants in active labor market programs has only a modest effect on elasticity estimates. Finally, the negatively sloping pay-unemployment relationship is only detected with overall regional unemployment rates, whereas education level-specific regional unemployment rates yield a statistically insignificant elasticity estimate close to zero.

Keywords: unemployment measurement, regional unemployment, earnings, wage curve

JEL Classification: J31, J64, R23

---

* The author thanks Jaakko Pehkonen and participants at the ERSA Congress (St. Petersburg) for helpful comments. The author gratefully acknowledges the financial support provided by the Finnish Cultural Foundation (grant 00120284).

1 Introduction

Starting from the ground-breaking work by Blanchflower and Oswald (1990, 1994), an extensive body of empirical evidence has confirmed the so called wage curve hypothesis stating that wages are lower in localities with higher unemployment. The wage curve relationship is commonly measured using the local unemployment elasticity of pay. Estimate of the wage curve elasticity is of considerable economic interest because it measures the degree of wage flexibility with respect to local unemployment conditions. Later empirical work has revealed that the elasticity estimates are not as uniform as suggested by the early findings: while Blanchflower and Oswald found that wages responded to local unemployment conditions in a remarkably similar way in countries with distinctly different labor market institutions – with a 100% increase in local unemployment rate being associated with a 10% decrease in wage level – later studies have documented substantial variation in wage curve estimates between countries (see, e.g., Nijkamp and Poot 2005). For example, as might be expected, wage responsiveness to local unemployment seems to be weaker in countries with centralized wage-setting systems (e.g., Nordic countries) than in countries with less centralized wage bargaining (e.g., the US).

Besides variation stemming from actual differences in labor market-related factors (e.g., wage-setting institutions), wage curve estimates are sensitive to methodological choices regarding, for example, which control variables are included and which estimation method is used in the analysis (e.g., Card 1995; Nijkamp and Poot 2005). One important choice is that of the unemployment measure: there are alternative ways to define unemployment, and it is not obvious which definition best captures the wage curve relationship. Theoretical explanations proposed for the wage curve relationship provide some guidance for the choice of the unemployment measure. These explanations often interpret regional unemployment rate as an indicator of local job competition, with increasing local unemployment tightening the competition for available jobs. The basic tenet of these explanations is that the inverse wage-unemployment relationship arises because the increasing job competition induced by the higher local unemployment worsens the outside options of the employees, which, in turn, weakens the bargaining power of employees (union bargaining model), reduces the wage levels needed to deter workers from shirking (efficiency wage model) and lowers the wage levels that employers are willing to pay in order to discourage workers from resigning (efficiency wage model with labor turnover); as a result, lower wages are paid in high-unemployment regions than in low-unemployment regions (Blanchflower and Oswald 1994; Campbell and Orszag 1998). However, the regional unemployment rate is a valid indicator of local job competition only if unemployed individuals actually generate competition for jobs by actively searching for work. For this to be true, only those nonemployed individuals who engage in active job search should be considered as unemployed.

In reality, local wage setting may not only depend on the number of nonemployed actively looking for work in the locality, but on a wider incidence of local joblessness – that is, it is possible that also less active nonemployed jobseekers exert downward pressure on local wage level. More generally, the ongoing dispute over the correct definition of unemployment (e.g., Jones and Riddell 1999, 2006) and the empirical evidence on measurement errors in unemployment rates (Feng and Hu 2013) highlight the importance of studying robustness of the wage curve elasticities to alternative measures of unemployment.

Given the extensive research on wage curve relationship, surprisingly few studies have examined the sensitivity of wage curve estimates to alternative definitions of unemployment. Pekkarinen (2001) studied the wage curve of metalworkers in Finland and found that adjusting standard unemployment rates to exclude long-term unemployed and to include participants in the subsidized employment schemes had only a modest effect on the estimated wage curve elasticities. With the exception of the study by Pekkarinen (2001), other prior studies on the topic have mainly focused on developing countries (Wu 2004; Kingdon and Knight 2006; Ilkkaracan et al. 2013), arguing that due to definition problems and idiosyncratic features of the labor markets, official unemployment statistics may provide an imperfect indicator of regional job search activity in less-developed countries. Correspondingly, Wu (2004) using data from China and Kingdon and Knight (2006) using data from South Africa detected the negatively-sloping wage curve with broader measures of unemployment (including, for example, the marginally-attached or laid-off workers), but not with official unemployment rates. Ilkkaracan et al. (2013), using data from Turkey, found that standard unemployment rates and broader measures of unemployment – including discouraged and marginally-attached workers – resulted in very similar wage curve estimates. Finally, Pannenberg and Schwarze (1998) examined the wage curve relationship in a transition economy, namely, in the regions of eastern Germany in the years following the German reunification, and argued that participants in active labor market programs should effectively be counted as unemployed and should hence be accounted for in the calculation of unemployment rates. Their findings illustrated the importance of using such modified unemployment rates while estimating the wage curve relationship, especially in economies with large-scale labor market programs: only the modified unemployment rates that incorporated the program participants yielded statistically significant negative pay-unemployment relationship, whereas standard unemployment rates yielded nonsignificant relationship.

Another issue related to the choice of regional unemployment variable when estimating the wage curve relationship is the relevant definition of 'local labor market'. If the unemployment rate of the local labor market is interpreted as an indicator of the level of job competition that workers face in the locality, one apparent way to determine a local labor market is to use information on spatial job search patterns: because the level of job competition depends on the number of individuals who generate competition by actively applying for open vacancies, the effective size of the local labor market can be determined based on the average distance from which firms with open vacancies in the locality receive job applications. In other words, the effective size of the local labor market depends on the distance at which the unemployed typically search for jobs from their places of residence (see, e.g., Manning and Petrongolo 2011). However, because detailed data on spatial job search patterns are generally unavailable, the wage curve relationship is typically estimated using unemployment data for administrative regions. A geographical disaggregation that is created for administrative purposes may yield regions that are effectively too small or too large to characterize functional local labor markets. Furthermore, functional local labor markets may overlap the borders of the administrative regions if there are substantial job search or commuting flows across neighboring regions.

Finally, when the regional unemployment rate is used as a proxy variable for the job competition faced by a worker, the regional unemployment among individuals with job-related attributes (e.g., education level or field of study) similar to his or her own may provide a more relevant measure than the overall regional unemployment. In the case of wage curve hypothesis, this would indicate that the wages of the individuals of a group with shared attributes are negatively related to the level of regional joblessness within that group, but are less related to the joblessness within other groups with distinct attributes. For example, a high regional unemployment among individuals with low education level may not exert downward pressure on wages of highly educated employees working in the same region.

In this study, we use micro panel data from Finland to examine the sensitivity of the wage curve elasticity to different measures (definitions) of unemployment and to different levels of geographical aggregation at which the local unemployment is measured. We first test the robustness of the wage curve relationship to different unemployment statistics, namely, the unemployment rates calculated based on the numbers of 1) the respondents of the Labour Force Survey who are classified as unemployed under the ILO (International Labour Organisation) definition of unemployment and 2) the individuals who have registered as unemployed at the employment offices. Furthermore, we examine the role that distinctive groups of nonemployed labor-force participants play in local wage-setting by estimating the wage curve with modified unemployment rates that exclude long-term unemployed and include participants in subsidized employment schemes. Next, to study whether the change in the level of regional disaggregation has an effect on the estimate of the local unemployment elasticity of pay, we estimate the wage curve relationship using unemployment rates measured at different regional levels. Finally, to assess the sensitivity of elasticity estimates to disaggregation of regional unemployment rates by relevant worker characteristics, we estimate the wage curve relationship using education level-specific regional unemployment rates.

2 Data and empirical approach

The data analyzed in this study were obtained from the population registers of Statistics Finland. The data are based on a 7% random sample of the Finnish population drawn in 2001, which includes relevant information on the sampled individuals for the 1995–2006 period. In our analysis, we focus on non-agricultural private sector employees who lived in mainland Finland[1]. In order to evaluate the relationship between pay level and local unemployment conditions, we estimate the following earnings equation:

$$log(e_{irt}) = a + \beta\, log(u_{rt}) + \boldsymbol{\gamma X_{irt}} + \delta_r + \theta_t + \varepsilon_{irt} \qquad (1)$$

where e_{irt} is the annual earnings for worker i in region r in year t; u_{rt} is a regional unemployment rate; $\boldsymbol{X_{irt}}$ is a vector of worker characteristics including age and work experience (and their squared terms), gender, native/first language, marital status (and its interaction with gender), dummy for children under 18 years old (and its interaction with gender), education level, field of study and industry; and δ_r and θ_t are region and year dummies, respectively[2]. The coefficient of interest is β, which measures the unemployment elasticity of pay. Based on the wage curve hypothesis, β is expected to be negative.

The dependent variable of the earnings equation is a logarithm of worker's annual earnings. According to the wage curve hypothesis, local unemployment rate is negatively related to *wages*, and hence the preferable dependent variable would be an hourly wage of a worker. Unfortunately, the data set used for the analysis does not include information on hourly wages or working hours. When annual earnings are used as a dependent variable, coefficient β may yield inflated estimates of wage-unemployment-relationship as β reflects variation in both *hourly wages* and *annual working hours* with respect to local unemployment level[3]. In order to account for the effects of varying working hours (and extreme earnings values) on the wage curve estimates, the sample used for the analysis excludes 1) employees who worked less than twelve months a year and 2) five percent of the observations from both tails of the earnings distribution in each year at each region. The final sample consists of 688,049 observations for 113,017 workers. However, as noted by Card (1995), the relevant dimension for the estimation of the unemployment coefficient β is not the number

1 Consequently, the final sample excludes public sector workers, individuals who lived in Åland Islands (which constitute an autonomous province of Finland) and workers who were employed by the agriculture, forestry or fishing industries. Additionally, employees aged under 18 were excluded from the sample.

2 The variables are described in more detail in Table A1 in the Appendix.

3 Previous studies demonstrate that wage curve estimates are inflated when annual earnings are used as a dependent variable instead of hourly wage (Card 1995; Bratsberg and Turunen 1996). For example, Bratsberg and Turunen (1996) found that the wage curve estimates decreased by about 40–50% when the logarithm of hourly wages was used as the dependent variable instead of the logarithm of annual earnings.

of worker-year observations but the product of the number of regions and the number of observation years. Consequently, when unemployment rates for NUTS-3 regions are used to estimate the earnings equation (1), the estimation of β is effectively based on 228 observations (= 19 regions*12 years).

In the baseline specifications, earnings equation (1) is estimated using an OLS estimator. Additionally, to account for the earnings effects of unobserved time-invariant worker characteristics, earnings equation is also estimated using the fixed-effects estimator (i.e. with a worker-specific effect η_i added to the right-hand side). Earnings equations including the worker fixed effects are our preferred specifications as they account for the potential composition bias that arises from the compositional changes in the workforce over the business cycle (Solon et al. 1994). To account for the possibility that the different level of aggregation of the dependent and the independent variable – earnings are measured at the worker-level and unemployment at the regional level – may cause error terms ε_{irt} to be correlate across employees working in the same region (Moulton 1986, 1990), we cluster standard errors at the regional level (see Cameron and Miller 2011, 2015).

To test the robustness of the wage curve coefficient β to different unemployment measures, we first estimate earnings equation (1) using four alternative definitions of unemployment (the definitions used are described in more detail in the next section). Second, to examine whether the estimate of β varies depending on the level of regional disaggregation used to measure the local unemployment level, we employ unemployment rates u_{rt} measured both at NUTS-3 level (19 regions) and at LAU-1 level (79 sub-regions) to estimate the earnings equation. Finally, to assess the importance of using regional unemployment data disaggregated by relevant worker characteristics for analyzing the local unemployment elasticity of pay, we utilize education level-specific regional unemployment rates u_{ert} (where e denotes the education level) to estimate the wage curve coefficient β.

Alternative unemployment measures

To test the robustness of the wage curve estimates to alternative unemployment measures, we first estimate the earnings equation using official regional unemployment rates drawn from two different data sources: *Labour Force Survey* data and *Employment Service Statistics* data on the unemployed registered at the employment offices[4]. Empirical research often uses these unemployment statistics interchangeably, which ignores the fact that due to differences in definitions, they measure different aspects of unemployment conditions. The unemployment rates drawn from the Labour Force Survey data ('survey-based unemployment') rely on the ILO definition of unemployment, according to which an unemployed is a person who is out of work, has actively searched for a job dur-

[4] The unemployment statistics based on the Labour Force Survey are provided by Statistics Finland. The Employment Service Statistics are compiled by the Ministry of Economic Affairs and Employment.

ing the past four weeks and is available to begin a job within the next two weeks. Conversely, the unemployment rates drawn from the Employment Service Statistics data ('register-based unemployment') include those nonemployed who are available for work, are seeking a full-time job (i.e. a job with more than 50% of regular working hours) and have officially registered as jobseekers.

The essential difference between the two definitions lies in the requirement for an active job search. Survey unemployment emphasizes the search activity because individuals are only considered unemployed if they have actively searched for a job recently. However, the activities considered as job search vary greatly; for example, asking among friends and placing job advertisements are both regarded as active job search (most importantly, registration at the employment office is not needed). In the case of registered unemployment, the requirement for active search is less stringent; the unemployed are obligated to contact the employment office at regular intervals after being registered as job applicants, but otherwise their personal job search activity can vary substantially.

Both the survey-based and the register-based unemployment rates have their shortcomings as measures of the prevailing unemployment level. Register-based unemployment statistics can be distorted by the inclusion of many individuals who are not actively searching for work and/or are not actually available for work (e.g., individuals involved in retirement schemes). Survey-based unemployment rates, on the other hand, may be mismeasured due to the misclassification of individuals' labour force statuses that results from errors in self-reported information (Feng and Hu 2013). Econometrically, such measurement errors in unemployment rates may cause wage curve estimates to be biased towards zero (Johnston and DiNardo 1997).

The two alternative unemployment rates for NUTS-3 regions are summarized in Table 1. The unemployment figures illustrate that the regional differences in unemployment were substantial; for example, in 2006 the survey-based regional unemployment rates ranged from 3.6% to 17.1%. Furthermore, the table shows that unemployment rates were very high at the beginning of the study period and decreased significantly later on. The initial high levels of unemployment are attributable to the severe recession Finland experienced in the early years of the 1990s. The recession led to a sharp increase of the regional unemployment rates, with some regions suffering more than the others: in 1990, the year before the economic crisis began, regional unemployment rates for NUTS-3 regions varied between 1% and 6%; by 1994, regional unemployment rates had increased to 13–22%.[5] Another observation from Table 1 is that the register-based unemployment rates regularly exceed those calculated from the survey data; the mean difference is 3.3 percentage points. The high levels of reg-

5 The unemployment figures reported in Table 1 are based on the Labour Force Survey, and exclude unemployment rates for Åland Islands. For a more detailed discussion of the recession in Finland in the early 1990s (often referred to as the Finnish Great Depression), see Gorodnichenko et al. (2012) and Gulan et al. (2014).

ister-based unemployment are primarily attributable to the fact that in Finland, unemployed must register at the employment offices to be eligible for unemployment benefits (see, e.g., Melis and Lüdeke 2006).

TABLE 1 Survey- and register-based unemployment rates (NUTS-3 regions)

	Source of the unemployment data					
	Survey	Register	Survey	Register	Survey	Register
Unemployment rate	1995–2006		1995		2006	
Mean U_r	11.7	15.0	16.5	20.3	8.8	10.8
Median U_r	11.4	14.5	16.5	20.6	8.8	11.4
Minimum U_r	3.6	6.6	11.7	14.9	3.6	6.6
Maximum U_r	23.5	26.0	22.4	25.7	17.1	16.4
Std. Dev. of U_r	3.9	4.4	2.9	3.0	3.0	2.9
$\text{Corr}(U_{r,\text{register}}, U_{r,\text{survey}})$	0.95		0.96		0.90	
$\text{Mean}(U_{r,\text{register}} - U_{r,\text{survey}})$	3.3		3.8		2.0	
$U_{r,\text{register}} > U_{r,\text{survey}}$	227/228					
$\text{Mean}(U_{r,2006} - U_{r,1995})$	−7.7	−9.5				

Notes: Subscript *r* refers to NUTS-3 region. "Survey" refers to unemployment rates computed from the Labour Force Survey data. "Register" refers to unemployment rates computed from the Employment Service Statistics.

In addition to survey- and register-based unemployment rates, we also employ two additional unemployment measures that were computed using the data drawn from the Employment Service Statistics. These unemployment measures are modified versions of the register-based unemployment rates. Based on the prior suggestions that unemployed who have been without work for a long period of time may not exert downward pressure on wages (e.g, Blackaby and Manning 1990), the first modified measure excludes the long-term unemployed – that is, unemployed who have been without a job for at least one year – from the standard register-based unemployment rates. The second measure is similar to that used by Pekkarinen (2001), and modifies standard unemployment rates in two ways: 1) long-term unemployed are excluded and 2) individuals participating in subsidized employment schemes are included[6]. As discussed and

[6] Pekkarinen (2001) analyzed the wage curve relationship for the Finnish metal industry workers in 1991–1995. Our analysis differs from that of Pekkarinen in two respects. First, our analysis is based on data on employees working in several different industries. Second, the data used by Pekkarinen includes the severe recession period of 1990–1993, while our data are from a period of uninterrupted economic growth (1995–2006). This may be important, as Boushey (2002) and Maczulskij (2013) provide

empirically illustrated by Pannenberg and Schwarze (1998) and Pekkarinen (2001), participants in active labor market programs should effectively be considered as unemployed job searchers in the wage curve analysis as local wage level is negatively related to the number of these participants in the region. Ideally, we would also include participants in other active labor market programs (e.g., participants in internships and job market training) as effective job searchers when computing the modified unemployment rates; unfortunately, however, our data include information only on participants in subsidized employment schemes. The exact formulas used to calculate the modified unemployment rates are presented in Table 2, which also reports descriptive statistics for all four regional unemployment measures used in the analysis.

TABLE 2 Alternative unemployment rates (NUTS-3 regions, 1995–2006)

Unemployment rate	Definition and source	Mean (s.d.)	Min Max
Survey-based	ILO definition of unemployment Labour Force Survey Statistics Finland	11.7 (3.9)	3.6 23.5
Register-based	Unemployed registered at the employment offices Employment Service Statistics Ministry of Economic Affairs and Employment	15.0 (4.4)	6.6 26.0
Register-based (modified 1)	$\frac{\text{Unemployed} - \text{long-term unemployed}}{\text{Labor force} - \text{long-term unemployed}}$	11.5 (3.6)	4.6 21.7
Register-based (modified 2)	$\frac{\text{Unemployed} - \text{long-term unemployed} + \text{individuals in subsidized employment schemes}}{\text{Labor force} - \text{long-term unemployed}}$	13.8 (4.7)	5.5 27.6

tentative evidence to suggest that the regional unemployment elasticity of pay varies over the business cycle.

3 Results

Table 3 reports the estimates of the unemployment elasticity of pay, β, based on the four alternative unemployment measures described in the previous section[7]. The results show that the size of the elasticity estimate is highly sensitive to the estimation method and the unemployment measure used. The elasticity estimates from the OLS regressions are statistically insignificant and close to zero. After controlling for worker fixed effects, an inverse pay-unemployment relationship is detected, but only when register-based unemployment rates are used. Hence, while some studies find that inclusion of worker fixed effects has merely a negligible effect on the estimate of unemployment elasticity of pay (e.g., Bratsberg and Turunen 1996; Sanz-de-Galdeano and Turunen 2006), our results illustrate that accounting for unobserved worker heterogeneity can be crucial for the detection of a statistically significant negative elasticity estimate. The specifications that include controls for worker fixed effects are the preferred ones, because they account for the potential composition bias that could obscure the true nature of the wage-unemployment-relationship (Solon et al. 1994). Hence, the following discussion focuses on the wage curve estimates obtained from the worker fixed effects models.

The results in columns (1) and (2) illustrate that alternative measures of regional unemployment can lead to notable differences in wage curve estimates: using register-based unemployment rates instead of survey-based unemployment rates increases the coefficient estimate on unemployment variable from a modest −0.012 to −0.064, and only the latter coefficient is statistically significant. Hence, the results imply that the "true" unemployment elasticity of pay may be obscured by an inappropriate choice of unemployment measure. Specifications in columns (3) and (4) are estimated using the modified unemployment rates. According to the results, the redefinition of register-based unemployment rates to exclude long-term unemployed (column 3) and to exclude long-term unemployed and to include participants in subsidized employment schemes (column 4) has only a negligible effect on wage curve estimates.

The finding that register-based unemployment yields a higher elasticity estimate than survey-based unemployment implies that local wage-setting decisions are contingent on a wider incidence of local unemployment, not simply on the number of nonemployed engaged in active job search. As a more general note, this finding suggests that in studies in which the local unemployment level is the primary predictor of interest, alternative unemployment measures should be used to assess the robustness of the results; otherwise, the true magnitude of the relationship between unemployment and the response variable in question may be obscured.

---

[7] For brevity, Table 3 only reports the unemployment coefficients. Table A2 in the Appendix reports coefficient estimates for other regressors of the specifications estimated in column (2).

TABLE 3 Wage curves based on alternative unemployment measures

Dependent variable: log(annual earnings)
Estimation period: 1995–2006
Regional level: NUTS-3 regions

	Unemployment rate			
	Survey-based	Register-based	Register-based (modified 1)	Register-based (modified 2)
OLS				
log(u_{rt})	0.007	−0.005	−0.001	−0.010
	(0.012)	(0.016)	(0.017)	(0.022)
Worker fixed effects				
log(u_{rt})	−0.012	−0.064	−0.058	−0.070
	(0.008)	(0.011)***	(0.011)***	(0.013)***
Worker characteristics	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
OLS				
$R^2_{adjusted}$	0.46	0.46	0.46	0.46
Worker fixed effects				
R^2 (within)	0.48	0.48	0.48	0.48
R^2 (between)	0.32	0.34	0.33	0.34
R^2 (overall)	0.33	0.34	0.34	0.34
F-test (worker FE)	15.76	15.72	15.72	15.71
(*p*-value)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)

Notes: Cluster-robust standard errors in parentheses, clustered at the regional level. The number of worker-year-observations is 688,049 (on 113,017 workers) and the number of region-year-observations is 228 (= 19 regions*12 years). Definitions of the alternative unemployment rates are presented in Table 2. Worker characteristics include the following variables: gender dummy, age, age^2, experience, experience^2, marital status dummy (and its interaction with gender), children dummy (and its interaction with gender), native/first language dummies, education level dummies, field of study dummies and industry dummies (gender and language dummies are omitted in the worker fixed effects specifications). Robust Hausman tests reject random effects in favor of fixed effects for all model specifications. The F-statistics test the joint statistical significance of the worker fixed effects ($F(113016, 574968)$ distribution; corresponding *p*-values in parentheses). The full results are available upon request. Significant at the * 10% level; ** 5% level; *** 1% level.

The elasticity estimate of −0.064 implies that on average, a 10% increase in the regional unemployment rate was associated with a 0.64% decrease in the local pay level. This estimate provides strong evidence in favor of the wage curve relationship, which predicts that pay level is lower in regions with higher unemployment. Hence, and in contrast to some earlier findings (Albæk et al. 2000), the results suggest that wages also responds to differences in the local unemployment conditions in countries such as Finland, where the regional wage adjustment is partly constrained by the highly centralized wage bargaining system. The observed estimate is close to the publication-bias-corrected estimate of −0.07 obtained by Nijkamp and Poot (2005) in their meta-analysis of wage curve studies from over 40 countries. Compared to prior studies of wage curve in Finland, our estimate indicates a somewhat more elastic wage curve than the ones observed by Pekkarinen (2001) for Finnish metal industry workers. Using register-based unemployment rates and model specifications similar to ours, Pekkarinen obtained wage curve estimates of −0.02 and −0.03. On the other hand, our wage curve estimate is smaller than those found by Maczulskij (2013), who observed wage curve coefficients ranging from −0.08 to −0.12 for Finnish private sector workers for the years 1990–2004.[8]

Next, to test how sensitive wage curve elasticity is to changes in the regional level at which the unemployment rate is measured, Table 4 provides separate wage curve estimates for 19 NUTS-3 regions (column 1) and for 79 LAU-1 sub-regions (column 2)[9]. The NUTS-3 regions are evidently too large to characterize functional local labor markets. The borders of the LAU-1 sub-regions, on the other hand, are in part determined to minimize the across border commuting flows, and are typically considered as reasonable approximations of local labor markets. The coefficient estimates in columns (1) and (2) are almost identical, suggesting that the level of geographical disaggregation at which the unemployment is measured has only a negligible influence on the elasticity of the wage curve. In contrast to the results obtained in Table 3, all OLS estimates are now negative, and with one exception, statistically significant at the 5% level. However, the worker fixed effects specifications yield substantially larger wage curve estimates than OLS specifications: in all columns, estimates differ by more than a factor of four. Again, the wage curve estimates

8 The difference between our estimates and the estimates in Maczulskij (2013) is apparently attributable to differences in the estimation approaches, particularly in sample restrictions. When models in Table 3 were estimated with a sample that did not impose any restrictions on the earnings distribution or on the number of working months, the OLS estimates inflated substantially (the estimates were typically close to −0.1) and became statistically significant, whereas the worker fixed effects estimates remained relatively unchanged (and hence did not alter the conclusions drawn here). This finding is expected since, as discussed above, wage curve estimates from annual earnings regressions are likely to be inflated if working hours are allowed to vary among workers.

9 The estimations in Table 4 are conducted for the period 1995–2002, because the region identifiers for LAU-1 regions were only available for these years. The unemployment rates for sub-regions are based on the numbers of registered unemployed (the survey-based unemployment rates are not available for LAU 1 sub-regions).

change only little when modified unemployment rates are used instead of the standard ones (columns 3 and 4).

TABLE 4 Wage curves based on alternative regional classifications

Dependent variable: log(annual earnings)
Estimation period: 1995–2002

Regional level	*NUTS-3*	*LAU-1*	*LAU-1*	*LAU-1*
Unemployment rate	Register-based	Register-based	Register-based (modified 1)	Register-based (modified 2)
OLS				
$\log(u_{rt})$	−0.016	−0.019	−0.021	−0.024
	(0.010)	(0.008)**	(0.009)**	(0.009)***
Worker fixed effects				
$\log(u_{rt})$	−0.093	−0.090	−0.092	−0.101
	(0.008)***	(0.012)***	(0.013)***	(0.013)***
Worker characteristics	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes
Worker-year observations	428,906	429,414	429,414	429,414
Regions	19	79	79	79
Observation years	8	8	8	8
Regions*years	152	632	632	632
OLS				
$R^2_{adjusted}$	0.43	0.44	0.44	0.44
Worker fixed effects				
R^2 (within)	0.40	0.40	0.40	0.40
R^2 (between)	0.31	0.31	0.31	0.32
R^2 (overall)	0.31	0.31	0.31	0.31
F-statistic	14.98	14.69	14.69	14.69
(*p*-value)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)

Notes: Cluster-robust standard errors in parentheses, clustered at the regional level. Definitions of the alternative unemployment rates are presented in Table 2. Worker characteristics include the following variables: gender dummy, age, age^2, experience, experience^2, marital status dummy (and its interaction with gender), children dummy (and its interaction with gender), native/first language dummies, education level dummies, field of study dummies and industry dummies (gender and language dummies are omitted in the worker fixed effects specifications). Robust Hausman tests reject random effects in favor of fixed effects for all model specifications. The *F*-statistics test the joint statistical significance of the worker fixed effects (NUTS-3 regions: $F(92660, 336186)$; LAU-1 regions: $F(92838, 336456)$; corresponding *p*-values in parentheses). The full results are available upon request. Significant at the * 10% level; ** 5% level; *** 1% level.

In the preceding analysis, local unemployment rate is treated as an exogenous variable. However, the wage curve estimates in Tables 3 and 4 may potentially suffer from endogeneity bias arising from the simultaneous determination of pay level and local unemployment: local unemployment level may have an effect on pay level (as predicted by the wage curve hypothesis) *and* local unemployment level may depend on the pay level in the locality. To account for the potential endogeneity of the unemployment variable, we followed several previous studies (e.g., Johnes 2007; Ilkkaracan et al. 2013) and instrumented the unemployment variable with its one-year lagged value. The instrumental variable (two-stage least squares) estimations yielded larger wage curve estimates than those reported above. For example, the coefficient estimates from the instrumental variable estimations of the worker fixed effects models of Tables 3 and 4 were, on average, 1.8 times larger than those reported in the tables. However, and more importantly, the instrumental variable estimations did not alter the main conclusions reached above[10].

Additional findings: Education level-specific unemployment rates

As discussed above, theoretical models proposed to explain the wage curve relationship typically interpret the overall local unemployment rate as a measure of prevailing job competition for individuals in the locality. However, local unemployment rates disaggregated by different demographic groups may provide a more appropriate measure for job competition than overall unemployment rates: because individuals with distinctive background characteristics generally compete for different jobs, the level of local job competition faced by a group of individuals who share some common characteristics (peer group) is essentially determined by the joblessness of the individuals in that particular group. Consequently, based on the wage curve hypothesis, the wages of the individuals within a peer group are likely to be negatively related to the group-specific unemployment rate, and less related to the unemployment rates of other groups with distinctive characteristics.

One relevant level of disaggregation is by education level. Following the reasoning above, the higher local unemployment of less educated individuals may not exert downward pressure on the wages of higher-educated workers, and vice versa. To assess the relationship between pay level and education level-specific unemployment level, we estimated earnings equation (1) using regional unemployment rates disaggregated into four education levels: 1) basic education, 2) upper secondary level education, 3) lowest level tertiary education/lower-degree level tertiary education and 4) higher-degree level tertiary

10 The full results of the instrumental variable estimations are available upon request. The use of lagged unemployment variable to instrument for unemployment variable is questionable, and, in the absence of more suitable instruments, these results are not discussed here in more detail.

education/doctorate or equivalent level education[11]. The unemployment rates varied substantially across education levels: the averages of regional education level-specific unemployment rates were 23%, 17%, 9% and 4%, respectively.

The results are presented in Table 5. The specifications in columns (1) and (2) are estimated using overall regional unemployment rate. As compared with the results in Table 3, these specifications are most comparable to those reported in the second column. The wage curve estimates observed here differ from those reported in Table 3, because the unemployment rates utilized in the estimations are computed using data from different sources[12]. Hence, these results once again illustrate that the slope of the wage curve depends on the unemployment measure used. The wage curve estimates in columns (3) and (4) are based on the education level-specific regional unemployment rates. The coefficient estimates from both OLS and worker fixed effects specifications indicate that education-level specific unemployment level in the locality played no role in the determination of pay level. To summarize, the results indicate that negatively sloping wage curve relationship is only detected with overall unemployment rates, but not with unemployment rates disaggregated by education level.

---

11 Education level-specific unemployment rates for NUTS-3 regions were computed using data from the *Employment statistics* of Statistics Finland. The unemployment figures reported in these statistics are based on the number of unemployed registered at the employment offices, and hence the computed unemployment rates correspond to register-based unemployment rates used in the analysis above. These statistics are based on a renewed regional classification, which slightly differs from that used in the preceding analysis: there are 18 (instead of 19) NUTS-3 regions, with two formerly distinct regions (namely, Uusimaa and Eastern Uusimaa) consolidated into one region. Labor market conditions in the consolidated regions were very similar during the analysis period: the averages of register-based unemployment rates for these regions were 10% (Uusimaa) and 9.9% (Eastern Uusimaa); the mean difference in unemployment rates was 0.5 percentage points. In our estimations, workers in both of these formerly distinct regions were matched with the unemployment rates of the consolidated NUTS-3 region.

12 Over the analysis period, the average difference between the overall regional unemployment rates used here and the unemployment rates used to estimate the specifications in the second column of Table 3 (excluding rates for Uusimaa and Itä-Uusimaa) was 0.65 percentage points.

TABLE 5 Wage curves based on education level-specific unemployment rates

Dependent variable: log(annual earnings)
Estimation period: 1995–2006
Regional level: NUTS-3 regions

	Overall regional unemployment		Education level-specific regional unemployment	
	OLS	Worker FE	OLS	Worker FE
$\log(u_{rt})$	0.006	−0.041		
	(0.016)	(0.011)***		
$\log(u_{ert})$			0.027	−0.001
			(0.018)	(0.013)
Worker fixed effects	No	Yes	No	Yes
Worker characteristics	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
R^2_{adjusted}	0.46		0.46	
R^2 (within)		0.48		0.48
R^2 (between)		0.33		0.32
R^2 (overall)		0.34		0.33
F-test (worker FE)		15.74		15.87
(*p*-value)		(< 0.001)		(< 0.001)

Notes: Cluster-robust standard errors in parentheses, clustered at the regional level. The number of worker-year-observations is 688,049 (on 113,017 workers) and the number of region-year-observations is 216 (= 18 regions*12 years). Unemployment rates were computed using data from the *Employment statistics* of Statistics Finland. $\log(u_{ert})$ is the logarithm of unemployment rate at education level *e* in region *r* in year *t*. Worker characteristics include the following variables: gender dummy, age, age^2, experience, experience^2, marital status dummy (and its interaction with gender), children dummy (and its interaction with gender), native/first language dummies, education level dummies, field of study dummies and industry dummies (gender and language dummies are omitted in the worker fixed effects specifications). Robust Hausman tests reject random effects in favor of fixed effects for specifications in columns 2 and 4. The *F*-statistics test the joint statistical significance of the worker fixed effects (*F*(*113016, 574969*)-distribution; corresponding *p*-values in parentheses). The full results are available upon request. Significant at the * 10% level; ** 5% level; *** 1% level.

4 Conclusions

Using worker-level panel data from Finland, this study examines the sensitivity of the slope estimates of the wage curve to changes in definitions of unemployment rate and local labor market. Three main conclusions can be drawn from the estimation results. First, the estimated wage curve slope varies considerably depending on the estimation method. The slope estimates from the standard OLS regressions are typically small and often statistically insignificant. The results of the worker fixed effects models provide strong evidence on the existence of wage curve relationship, as coefficient estimates generally indicate that workers in low-unemployment regions had higher pay levels than identical workers in high-unemployment regions.

Second, alternative measures of local unemployment produce different estimates of the wage curve slope. The wage curve estimates from a preferred model specification show that the unemployment rates computed from the Labour Force Survey data based on the ILO definition of unemployment provided substantially smaller wage curve estimate (−0.012) than the unemployment rates computed from the administrative data on registered unemployed (−0.064); furthermore, only the latter estimate was statistically significant. On the other hand, modifying standard unemployment rates to exclude long-term unemployed and to include participants in active labor market programs had only a modest effect on elasticity estimates.

Third, the wage curve estimate is not sensitive to changes in the level of geographical disaggregation at which the unemployment rates are measured: unemployment rates of the NUTS-3 regions (19 regions) and LAU-1 regions (79 regions) provide identical wage curve slopes. Fourth, the negatively sloping pay-unemployment relationship is only detected with overall regional unemployment rates, whereas education level-specific regional unemployment rates yield a statistically insignificant elasticity estimate close to zero.

References

Albæk, K., Asplund, R., Barth, E., Blomskog, S., Guômundsson, B.R., Karlsson, V. and Strøjer Madsen, E. 2000. Dimensions of the wage-unemployment relationship in the Nordic countries: Wage flexibility without wage curves. In: Polachek, S. and Tatsiramos, K. (eds.) *Research in Labor Economics*, Volume 19, pp. 345–381.

Blackaby, D.H. and D.N. Manning. 1990. The north-south divide: Questions of existence and stability. *The Economic Journal* 100(401), 510–527.

Blanchflower, D.G. and Oswald, A.J. 1990. The wage curve. *Scandinavian Journal of Economics* 92(2), 237–242.

Blanchflower, D. G. and Oswald, A.J. 1994. *The Wage Curve*. The MIT Press: Cambridge, MA, US.

Boushey, H. 2002. Reworking the wage curve: Exploring the consistency of the model across time, space and demographic group. *Review of Political Economy* 14(3), 293–311.

Bratsberg, B. and Turunen, J. 1996. Wage curve evidence from panel data. *Economics Letters* 51(3), 345–353.

Cameron, A.C. and Miller, D.L. 2011. Robust inference with clustered data. In: Ullah, A. and Giles, D.E.A. (eds.) *Handbook of Empirical Economics and Finance*, pp. 1–28. Boca Raton, FL, US: Chapman & Hall/CRC.

Cameron, A.C. and Miller, D.L. 2015. A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317–372.

Campbell, C. and Orszag, J.M. 1998. A model of the wage curve. *Economics Letters* 59(1), 119–125.

Card, D. 1995. The Wage Curve: A Review. *Journal of Economic Literature* 33(2), 785–799.

Feng, S. and Hu, Y. 2013. Misclassification errors and the underestimation of the US unemployment rate. *American Economic Review* 103(2), 1054–1070.

Gorodnichenko, Y., Mendoza, E.G. and Tesar, L.L. 2012. The Finnish great depression: From Russia with love. *American Economic Review* 102(4), 1619–1643.

Gulan, A., Haavio, M. and Kilponen, J. 2014. Kiss me deadly: From Finnish great depression to great recession. *Bank of Finland Research Discussion Papers* 24: 2014.

Ilkkaracan, I., Levent, H. and Polat, S. 2013. Exploring different measures of wage flexibility in a developing economy context: the case for Turkey. *International Review of Applied Economics* 27(3), 297–315.

Johnes, G. 2007. The wage curve revisited: Estimates from a UK panel. *Economics Letters* 94(3), 414–420.

Johnston, J. and DiNardo, J. 1997. *Econometric Methods*. New York, NY, US: McGraw-Hill.

Jones, S.R.G. and Riddell, W.G. 1999. The measurement of unemployment: An empirical approach. *Econometrica* 67(1), 147–162.

Jones, S.R.G. and Riddell, W.G. 2006. Unemployment and nonemployment: Heterogeneities in labor market states. *The Review of Economics and Statistics* 88(2), 314–323.

Kingdon, G.G. and Knight, J. 2006. How flexible are wages in response to local unemployment in South Africa?. *ILR Review* 59(3), 471–495.

Maczulskij, T. 2013. Public–private sector wage differentials and the business cycle. *Economic Systems* 37(2), 284–301.

Manning, A. and Petrongolo, B. 2011. How local are labor markets? Evidence from a spatial job search model. *IZA Discussion Paper Series* No. 6178.

Melis, A. and Lüdeke, B. 2006. Registered unemployment (RU) compared with harmonised unemployed (LFS). European Commission, *Working papers and studies*: Brussels, Belgium.

Moulton, B.R. 1986. Random group effects and the precision of regression estimates. *Journal of Econometrics* 32(3), 385–397.

Moulton, B.R. 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics* 72(2), 334–338.

Nijkamp, P. and Poot, J. 2005. The last word on the wage curve?. *Journal of Economic Surveys* 19(3), 421–450.

Pannenberg, M. and Schwarze, J. 1998. Labor market slack and the wage curve. *Economics Letters* 58(3), 351–354.

Pekkarinen, T. 2001. The wage curve: Evidence from the Finnish metal industry panel data. *Finnish Economic Papers* 14(1), 51–60.

Sanz-de-Galdeano, A. and Turunen, J. 2006. The euro area wage curve. *Economics Letters* 92(1), 93–98.

Solon, G., Barsky, R. and Parker, J.A. 1994. Measuring the cyclicality of real wages: How important is composition bias?. *The Quarterly Journal of Economics* 109(1), 1–25.

Wu, Z. 2004. Wage curve for urban China: A panel data approach. *Applied Economics Letters* 11(7), 425–428.

Appendix

TABLE A1 Description of the variables

Variable	Description
Annual earnings	Annual earnings in euros.
Regional unemployment rate	Regional unemployment rate. Depending on the estimated specification, the unemployment rate is computed either at NUTS-3 level (19 regions) or at LAU-1 level (79 sub-regions). The sources of the unemployment data: 1) Labour Force Survey (LFS), Statistics Finland, 2) Employment Service Statistics, Ministry of Economic Affairs and Employment
Gender	Female dummy: = 1 if female, = 0 if male
Age	Age in full years
Experience	Work experience in years (since 1987). The values are calculated as follows: (The number of working months since January 1987) / 12
Marital status	Marriage dummy: = 1 if married, = 0 otherwise
Native language	Dummy variables for native/first language: 1) Finnish, 2) Swedish or 3) other
Children	Children dummy: = 1 if the worker had children under 18 years, = 0 otherwise
Level of education (based on ISCED 1997)	1) Primary education or lower secondary education (or unknown), 2) Upper secondary level education, 3) Lowest level tertiary education, 4) Lower-degree level tertiary education, 5) Higher-degree level tertiary education, 6) Doctorate or equivalent level tertiary
Field of study (based on ISCED 1997 classification)	1) General programmes (or not known or unspecified), 2) Education, 3) Humanities and arts, 4) Social sciences, business and law, 5) Science, 6) Engineering, manufacturing and construction, 7) Agriculture, 8) Health and welfare, 9) Services
Industry (based on NACE classification)	1) Agriculture, forestry and fishing (excluded from the analysis), 2) Mining and quarrying, 3) Manufacturing, 4) Electricity, gas and water supply, 5) Construction, 6) Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods, 7) Hotels and restaurants, 8) Transport, storage and communication, 9) Financial intermediation, 10) Real estate, renting and business activities, 11) Public administration and defence; compulsory social security, 12) Education, 13) Health and social work, 14) Other community, social and personal service activities, 15) Private households employing domestic staff and undifferentiated production activities of households for own use; extra-territorial organizations and bodies; industry unknown

Notes: Source of education and industry classifications: Statistics Finland.

TABLE A2 Earnings equations

Dependent variable: log(annual earnings) Estimation period: 1995–2006 Regional level: NUTS-3 regions				
	OLS		Worker fixed effects	
	Coefficient	Robust SE	Coefficient	Robust SE
$\log(u_{rt})$	−0.005	(0.016)	−0.064	(0.011)***
Female	−0.170	(0.006)***		
Age	0.015	(0.005)***	0.030	(0.004)***
Age^2	−0.0002	(0.0001)***	−0.0004	(0.0001)***
Work experience	0.039	(0.002)***	0.079	(0.003)***
Work experience^2	−0.0007	(0.0001)***	−0.001	(0.00002)***
Married	0.043	(0.007)***	0.013	(0.002)***
Married*Female	−0.050	(0.007)***	−0.042	(0.003)***
Children dummy	0.004	(0.003)	−0.006	(0.004)
Children dummy*Female	−0.047	(0.007)***	−0.077	(0.012)***
<u>Education level</u>				
Primary/lower secondary	(*omitted*)		(*omitted*)	
Upper secondary	0.084	(0.017)***	−0.031	(0.011)**
Lowest level tertiary	0.214	(0.023)***	0.081	(0.025)***
Lower-degree level tertiary	0.343	(0.033)***	0.177	(0.016)***
Higher-degree level tertiary	0.506	(0.039)***	0.330	(0.015)***
Doctorate or equivalent	0.558	(0.034)***	0.457	(0.034)***
<u>Language dummies</u>				
Finnish	(*omitted*)			
Swedish	−0.025	(0.015)		
Other/unknown	0.017	(0.005)***		
<u>Other controls</u>				
Field of study	Yes		Yes	
Industry	Yes		Yes	
Year	Yes		Yes	
Region	Yes		Yes	
$R^2_{adjusted}$	0.46			
R^2 (within)			0.48	
R^2 (between)			0.34	
R^2 (overall)			0.34	

Notes: These are the same specifications estimated in Column (2) of Table 3. Robust SE = cluster-robust standard error (clustered at the regional level). The number of worker-year-observations is 688,049 (on 113,017 workers). Children dummy equals one if a worker had children under 18 years old and zero otherwise. Significant at the * 10% level; ** 5% level; *** 1% level.

SUMMARY IN FINNISH (YHTEENVETO)

Tutkimuksia palkoista, ylenemisistä ja suoriutumisarvioinneista

Väitöskirja koostuu neljästä empiirisestä tutkimusartikkelista. Kahdessa ensimmäisessä artikkelissa (luvut 2 ja 3) hyödynnetään erään suomalaisen yliopiston henkilöstöaineistoa vuosilta 2006–2012 tarkasteltaessa sitä, miten työntekijöiden suoriutumisarviointi ja palkitseminen riippuvat mitattavissa olevasta työsuoriutumisesta ja erilaisista taustatekijöistä. Henkilöstöaineisto sisältää ainutlaatuista ja yksityiskohtaista tietoa työntekijöiden taustaominaisuuksista, työssä suoriutumisesta ja työtehtävien vaativuudesta, ja antaa näin erinomaiset lähtökohdat arvioida niitä tekijöitä, jotka määrittävät esimiesten alaisistaan tekemiä suoriutumisarviointeja ja työntekijöiden ylenemistä vaativampiin työtehtäviin. Kahdessa jälkimmäisessä tutkimusartikkelissa (luvut 4 ja 5) tarkastellaan paikallisten työmarkkinaolosuhteiden merkitystä palkkojen määräytymisessä analysoimalla palkkatason ja alueellisen työttömyysasteen välistä yhteyttä. Näiden artikkelien tarkastelut perustuvat Tilastokeskuksen Suomen vuoden 2001 väestöstä tekemään seitsemän prosentin satunnaisotokseen. Tutkimusaineisto sisältää otokseen päätyneistä henkilöistä analyysien kannalta relevantit muuttujatiedot vuosille 1995–2006. Tähän yksilötason pitkittäisaineistoon on liitetty analyysejä varten alueellisia työttömyysastetietoja, jotka perustuvat Tilastokeskuksen työvoimatutkimukseen ja työ- ja elinkeinoministeriön työnvälitystilastoon.

Luvun 2 tutkimusartikkeli tarkastelee työntekijän mitattavissa olevien työtulosten ja työnantajakohtaisen työkokemuksen roolia esimiesten tekemissä suoriutumisarvioinneissa ja ylenemisissä vaativampiin työtehtäviin/korkeammille organisaatiotasoille. Aineiston kuvaileva analyysi paljastaa, että vaativammissa työtehtävissä työskennelleiden työntekijöiden suoriutuminen arvioitiin keskimäärin korkeammaksi kuin vähemmän vaativissa tehtävissä työskennelleiden, ja että suoriutumisarviot kasautuivat yhdeksänportaisella arviointiasteikolla tyypillisesti muutamiin ”normiarvoihin”. Lisäksi kuvaileva analyysi osoittaa, että työntekijöiden suoriutumisarvioiden alentaminen sekä työntekijöiden siirtäminen alhaisemman vaativuustason työtehtäviin oli hyvin harvinaista. Aineiston ekonometrisen analyysin tulokset osoittavat, että paremmin työssään suoriutuvat työntekijät – mitattaessa suoriutumista joko absoluuttisella työtuotoksella tai tuotoksella suhteessa kollegoihin – saivat muita parempia suoriutumisarvioita ja ylenivät muita todennäköisemmin vaativampiin työtehtäviin. Tulokset osoittavat myös, että pidempään saman työnantajan palveluksessa palvelleiden työntekijöiden työsuoriutuminen arvioitiin paremmaksi kuin heidän yhtä tuotteliaiden mutta vähemmän työkokemusta kerryttäneiden kollegoidensa suoriutuminen. Tulosten perusteella kokeneemmat työntekijät näyttivät myös ylenneen kokemattomampia kollegoitaan todennäköisemmin, joskaan tulokset eivät olleet näiltä osin täysin yhtenäisiä.

Luvun 3 artikkelissa analysoidaan työntekijöiden suoriutumisarvioiden, ylenemistodennäköisyyksien ja palkkatasojen vaihtelua sukupuolen perusteella.

Tulokset osoittavat, että vakioitaessa erot työntekijöiden taustaominaisuuksissa ja (julkaisuaktiivisuudella mitatussa) työtuotoksessa mies- ja naistyöntekijät saivat tyypillisesti yhtäläisiä suoriutumisarviointeja ja ylenivät yhtä todennäköisesti vaativampiin työtehtäviin. Tulokset myös havainnollistavat, että miesten ja naisten keskipalkoissa havaittava ero selittyi merkittäviltä osin eroilla työntekijöiden taustaominaisuuksissa ja työtuotoksessa. Julkaisuaktiivisuutta tarkastelevan lisäanalyysin tulosten perusteella naistutkijat tuottivat vähemmän vertaisarvioituja artikkeleita kuin taustaominaisuuksiltaan samankaltaiset miestutkijat.

Luvun 4 artikkelissa arvioidaan työmarkkina-alueella vallitsevan työttömyystilanteen roolia palkkatasoa määrittävänä tekijänä. Artikkelin tutkimuskohteena on ns. palkkakäyrärelaatio, jonka mukaan palkat ovat keskimäärin alhaisempia korkean työttömyyden alueilla, kun muut palkkatasoon vaikuttavat tekijät pidetään vakioituina. Artikkelissa tarkastelun kohteena on palkkakäyrärelaation alueellinen vaihtelu. Tutkimuksessa pyritään vastaamaan erityisesti seuraavaan kysymykseen: Onko palkkojen joustavuus paikallisen työttömyystilanteen suhteen voimakkaampaa syrjäisemillä alueilla ja alueilla joille on kasaantunut vähemmän taloudellista toimintaa? Tutkimuskysymyksen taustalla on hypoteesi, jonka mukaan tällaisten alueiden työnantajilla on enemmän monopsonivoimaa suhteessa työntekijöihinsä kuin kasvukeskittymissä toimivilla työnantajilla, ja tämä markkinavoima mahdollistaa palkkojen voimakkaamman sopeuttamisen paikallisen työmarkkinatilanteen heiketessä. Analyysin tulokset tuottavat vahvoja todisteita palkkakäyrärelaation tueksi niiden osoittaessa johdonmukaisesti, että palkat ovat Suomessa alhaisempia korkean työttömyyden alueilla. Toisin sanoen, palkkojen työttömyysjousto on Suomessa negatiivinen. Tulokset myös havainnollistavat, että palkkojen joustavuudessa aluetyöttömyyden suhteen esiintyy vaihtelua alueiden välillä (joskaan tämä vaihtelu ei ole kovin merkittävää). Tulokset eivät tuota todisteita monopsonivoimaan nojaavan hypoteesin tueksi: preferoidun, kiinteät työntekijävaikutukset huomioivan regressiomallin tulosten perusteella palkkojen työttömyysjousto ei riipu mittarista, joka huomioi alueen ”syrjäisyyttä” ja taloudellisen toiminnan kasaantumista.

Luvun 5 artikkelissa testataan, miten palkkakäyrärelaation voimakkuus vaihtelee sen estimoimisessa käytettävän työttömyysmuuttujan ja tilastollisen analyysimenetelmän perusteella. Tulokset korostavat analyysimenetelmän merkitystä palkkakäyrän havaitsemisessa: palkkojen ja aluetyöttömyyden välille havaitaan tilastollisesti merkitsevä käänteinen suhde ainoastaan käytettäessä ns. kiinteiden vaikutusten mallia, kun taas pienimmän neliösumman menetelmällä arvioituna muuttujien välillä ei näytä useimmiten olevan tilastollisesti merkitsevää yhteyttä. Tulokset myös osoittavat, että palkkojen työttömyysjoustolle havaittu estimaatti vaihtelee huomattavasti estimoinnissa käytetyn työttömyysmuuttujan perusteella, kun taas sillä mitä aluetasoa työttömyyden mittaamisessa käytetään ei näytä olevan suurta merkitystä estimaatin arvoon. Palkkakäyrärelaatiota ei havaita koulutustasoittain eritellyillä aluetyöttömyysasteilla, vaan ainoastaan alueellisilla kokonaistyöttömyysasteilla. Tämä tulos viittaa siihen, että palkkojen määräytymisen kannalta merkityksellistä on työssäkäyntialueen yleinen työttömyystilanne, ei niinkään työntekijäryhmäkohtainen työttömyystilanne.